Examining the Validity of Accelerometry for Measuring Arm Movements in Children

by

Jaclyn Kelly Dawe

A thesis submitted in conformity with the requirements for the degree of Master of Science in Rehabilitation Science
Rehabilitation Sciences Institute
University of Toronto

© Copyright by Jaclyn Kelly Dawe 2017
Examining the Validity of Accelerometry for Measuring Arm Movements in Children

Jaclyn Kelly Dawe

Master of Science in Rehabilitation Science
Rehabilitation Sciences Institute
University of Toronto
2017

Abstract

A goal of rehabilitation for children with hemiparesis is improved arm use during daily life. However, paediatric outcome measures do not objectively measure arm use in natural environments. Accelerometers, which quantitatively measure movement, may provide some missing information. This thesis examines the criterion, convergent and discriminative validity of accelerometry as a measure of paediatric arm movement. The results from a sample of 26 typically-developing (TD) participants and 13 participants with hemiplegic cerebral palsy (CP) indicate arm-accelerometry has: i) criterion validity, based on its sensitivity, specificity (Youden’s Index, \( J > .63 \) and agreement (> 84%), with reference to video-based observations; ii) moderate convergent validity, based on the correlation between arm ratios and QUEST scores \( (r_s = 0.662, p = .026) \); and iii) discriminative validity, based on differences in arm ratios \( (p < 0.02) \) between 9 matched pairs with and without CP. These findings warrant ongoing research on arm-accelerometry’s psychometric properties in children.
Acknowledgments

I am incredibly thankful to my supervisor, Dr. Kristin Musselman, for her guidance, mentorship, and sincere dedication to making the research process a valuable journey. In modeling integrity and applying her wisdom and vision to all her endeavours, she continues inspire me in pursuing research. I am also greatly thankful to my co-supervisor, Dr. José Zariffa, for accepting me as a student and introducing me to new areas of academia and research, with patient guidance and encouragement. Both of my supervisors’ responsiveness and willingness to share helpful insights through collaboration have allowed me to grow academically and professionally.

I would also like to thank my advisory committee, Dr. Darcy Fehlings, Dr. Peter Rumney, and Dr. Rosalie Wang. Their insightful advice and contributions were valuable in improving the quality and scope of this research project, and in enriching my learning. To Dr. Jaynie Yang, co-investigator, I am so thankful for her openness to collaboration and assistance with recruitment.

As well, I am sincerely thankful to the Director of the Rehabilitation Sciences Institute (RSI), Dr. Angela Colantonio, and the other professors who taught my graduate courses. They inspired me by providing intellectual challenges with openness and support for student-led initiatives.

To the graduate students in my cohort, I am thankful for their genuine peer support, in helping to create a collaborative working environment and graduate school journey. To Jirapat Likitlersuang, Vera Zivanovic and Nikola Unic, I am so grateful for their research assistance.

To all participants and families who volunteered their time to this project, I owe great thanks. It was an honour to work with them in a research capacity, and I wish them the best in the future.

Finally, I owe my greatest thanks to my parents, whose unconditional love and support has allowed me to be here in Toronto, pursuing this academic path. Their true presence in my life continues to uplift my spirits throughout each challenge that I face.

I would also like to acknowledge Toronto Rehabilitation Institute, for funding of this project.
# Table of Contents

Acknowledgments .......................................................................................................................... iii  
Table of Contents .......................................................................................................................... iv  
List of Tables ................................................................................................................................... viii  
List of Figures ............................................................................................................................... ix  
List of Appendices .......................................................................................................................... x  

Chapter 1 Introduction .................................................................................................................. 1  
1.1 Rationale for Study .................................................................................................................. 1  
1.2 Objectives of Study .................................................................................................................. 2  
1.3 Potential Impact of this Study ................................................................................................. 4  

Chapter 2 Literature Review ......................................................................................................... 5  
2 Hemiparesis ................................................................................................................................. 5  
2.1 Prevalence and Significance of Childhood Hemiparesis ......................................................... 5  
2.2 Therapeutic Interventions for Childhood Hemiparesis .......................................................... 7  
2.3 Treatment Outcome Measures for Childhood Hemiparesis .................................................. 8  
2.4 Accelerometry ........................................................................................................................ 13  
2.5 Validation of Accelerometry as a Measure of Arm Use ......................................................... 14  
2.6 The Need for an Objective Measure of Daily Arm Movement in Children .......................... 16  
2.7 Validity and its Relevance to this Study .................................................................................. 16  

Chapter 3 Methodology ................................................................................................................ 18  
3.1 Overview .................................................................................................................................. 18  
3.2 Participants ............................................................................................................................... 18  
3.3 Recruitment of Participants ..................................................................................................... 20  
3.4 Study Procedures ...................................................................................................................... 21
3.5 Data Analysis .......................................................................................................................... 26
  3.5.1 Analysis of Videotape Data ................................................................................................. 26
  3.5.2 Inter-Rater Reliability Testing ............................................................................................ 26
  3.5.3 Arm Accelerometry Data Processing .................................................................................. 28
  3.5.4 Receiver Operating Characteristic (ROC) Analysis .......................................................... 29
  3.5.5 Accelerometry-based Arm Movement Ratios .................................................................... 30
  3.5.6 Evaluation of Criterion Validity ......................................................................................... 30
  3.5.7 Disregard Index Calculation for Participants with Hemiparesis ...................................... 31
  3.5.8 Evaluation of Convergent Validity ..................................................................................... 32
  3.5.9 Evaluation of Discriminative Validity ................................................................................ 33

Chapter 4 ................................................................................................................................... 35

4 Results ..................................................................................................................................... 35

  4.1 Participant Demographics ..................................................................................................... 35
    4.1.1 Typically-developing (TD) Group .................................................................................... 35
    4.1.2 Participants with Hemiparesis – CP Group ................................................................. 35
  4.2 Inter-Rater Reliability of Video Observations ....................................................................... 37
    4.2.1 PABAK (Nominal Scores) ............................................................................................. 37
    4.2.2 PABAK (Ordinal Scores) .............................................................................................. 37
  4.3 Criterion Validity of Accelerometry - Typically-developing (TD) Group ......................... 37
    4.3.1 ROC Analysis: Determining an Optimal Threshold Setting ........................................ 38
    4.3.2 Sensitivity, Specificity and Youden’s Index at Different Thresholds ........................... 39
    4.3.3 Observer vs. Accelerometer Agreement Levels at Different Thresholds .................. 40
    4.3.4 PABAK Statistic – Accelerometer-Observer Agreement at Different Thresholds .......... 40
  4.4 Arm Use Ratios in Typically-developing Group .................................................................. 41
    4.4.1 Arm Use Ratios – Comparison between age groups .................................................. 41
4.5 Results for Group with Hemiparesis ................................................................. 42
  4.5.1 Descriptive and Clinical Measures for the Group with Hemiparesis ............... 42
4.6 Accelerometer Accuracy – Group with Hemiparetic Cerebral Palsy (CP) ............. 44
  4.6.1 ROC Analysis: Sensitivity and Specificity at Different Threshold Settings ....... 44
  4.6.2 Threshold for Optimal Sensitivity and Specificity ........................................ 45
  4.6.3 Sensitivity, Specificity and Youden’s Index at Different Thresholds .............. 45
  4.6.4 Observer vs. Accelerometer Agreement at Different Thresholds – CP Group .... 46
4.7 Correlational Analysis: Clinical Scores vs. Arm Use Ratios – CP Group .............. 48
  4.7.1 QUEST Scores vs. Accelerometry-based Arm Use Ratios ............................. 48
  4.7.2 PMAL-R Scores vs. Accelerometry-based Arm Ratios ................................. 51
  4.7.3 Additional Exploratory Analyses ................................................................. 51
4.8 Arm Movement Ratios: TD Group vs. Group with CP ...................................... 52
  4.8.1 Arm Movement Ratios At an Accelerometry Threshold of 2.0 ...................... 52
  4.8.2 Arm Movement Ratios At a Threshold Excluding Very Small Movements ...... 53
5 Discussion .............................................................................................................. 55
  5.1 Inter-rater reliability testing ............................................................................. 55
  5.2 Determining an Optimal Threshold Setting for Accelerometer Vector Magnitudes .... 56
  5.3 Accelerometer Accuracy .................................................................................. 57
  5.4 Accelerometry-based Arm Movement Ratios .................................................. 59
  5.5 Rationale for the Decision to Eliminate Very Small Movements ...................... 61
  5.6 Convergent Validity Analysis ......................................................................... 62
  5.7 Discriminative Validity Testing ....................................................................... 64
  5.8 Limitations of Study ....................................................................................... 64
  5.9 Implications of Research Findings: .................................................................. 66
Chapter 6........................................................................................................................................67

6 Conclusions..................................................................................................................................67

6.1 Concluding Summary ..................................................................................................................67

Chapter 7........................................................................................................................................69

7 Future Directions..........................................................................................................................69

7.1 Directions for Future Research ..................................................................................................69

References.........................................................................................................................................71

Appendices.......................................................................................................................................80
List of Tables

Table 1 Classification of Paediatric Outcome Measures for Hemiparetic Upper Extremities…… 5
Table 2 Interpretation of PABAK Scores .................................................................................. 23
Table 3 Demographic Characteristics of Typically-Developing (TD) Group............................. 30
Table 4 Demographic Characteristics of Children with Hemiparesis (CP Group)...................... 31
Table 5 Descriptive and Clinical Measures for Participants with Hemiparesis ......................... 38
Table 6 Mean Accelerometry-based Arm Ratios for Matched Pairs ........................................... 49
List of Figures

Figure 1 Structure of Play Time.................................................................17
Figure 2 Frequency Distribution of Accelerometry-Based Optimal Thresholds-TD Group........ 33
Figure 3 Sensitivity, Specificity and Youden’s Index Values-TD Group......................................................34
Figure 4 Comparing Agreement Levels at Thresholds of Zero, 7.1 and 2.0..........................35
Figure 5 PABAK at Thresholds of Zero, 7.1 and 2.0 .........................................................36
Figure 6 Mean Arm Movement Ratios for TD Group Participants, by Age Group................37
Figure 7 Frequency Distribution of Accelerometry-Based Optimal Thresholds-CP Group........40
Figure 8 Sensitivity, Specificity and Youden’s Index Values-CP Group.................................41
Figure 9 Comparing Agreement Levels at Thresholds of Zero, 4.88 and 2.0-CP Group ...........42
Figure 10 PABAK at Thresholds of Zero, 4.88 and 2.0-CP Group...........................................43
Figure 11 Unilateral QUEST Scores and Arm Use Ratios .......................................................45
Figure 12 PMAL-R(HO) Scores and Arm Use Ratios for Participants with CP ......................46
Figure 13 Arm Ratios of Matched Pairs from TD and CP Groups...........................................47
Figure 14 Arm Ratios of Matched Pairs from CP and TD Groups............................................48
List of Appendices

Appendix A Data Collection Sheet ........................................................................................................... 76
Chapter 1 Introduction

The current study involved an examination of the validity of arm accelerometry to measure arm movement in children. Accelerations of the arms, which indicated the presence of movement, were examined for both typically-developing (TD) children and children with hemiparesis (due to cerebral palsy (CP)). The primary aim of the study was to examine accelerometry’s criterion validity, by comparing its detection of arm use to concurrently gathered video-based observations of arm use, during each child’s engagement in functional play activities, in a lab setting. Secondly, this tool’s convergent validity was examined by comparing its quantification of relative affected arm use in children with hemiparesis to their scores on standardized clinical measures of bilateral and unilateral arm function. Finally, this tool’s ability to distinguish between children with hemiparesis (due to CP) and age- and sex-matched TD children was tested by comparing accelerometry-based measures of relative bilateral arm use (i.e. arm use ratios) between these two groups (i.e. TD and CP groups).

1.1 Rationale for Study

Hemiparesis affects one-third of children with CP, which is the most common pediatric physical disability, affecting over 500,000 children and families across North America. Although childhood stroke is much rarer than CP, a great proportion of children with post-natal stroke develop hemiparesis as well. Children with hemiparesis often present with ‘developmental disregard’ of the affected upper limb. This is characterized by levels of affected arm use that are even lower than those expected based on their physical capacity. Therapeutic interventions for these children often target increased functional use of the affected arm in daily life. However, in evaluating treatment outcomes, most standardized pediatric measures involve eliciting the child’s optimal capacity for arm use, rather than measuring actual daily arm use. There are consistently reported discrepancies between performance on these tests and children’s continued use of the affected arm in daily life. Such discrepancies have motivated the development of measures targeting children’s arm use in natural environments. However, currently available tools all rely on subjective parent/child report, which is prone to recall bias and varying standards between interviewees. There is still no validated objective measure of a child’s affected arm use in natural environments. Arm accelerometry can potentially be used to measure arm/hand use in natural environments, based on objective quantification. In this respect, it could be valuable in providing
information that is missing from both clinic-based child assessments and subjective report measures of pediatric arm use.

Although arm accelerometry has not yet been validated for this purpose in children, its validity as a measure of arm use in adults with hemiparesis has been established. Uswatte et al. (2000) first provided evidence of its criterion validity, comparing accelerometry-based arm use scores to videotaped observations of movement. Since that time, several studies, including two systematic reviews, have further supported the validity of accelerometry-based arm ratios as an index of real-world rehabilitative outcome in post-stroke adults with hemiparesis. However, only one study (Sokal et al., 2015) has explored this use of arm accelerometry in children with hemiparesis, finding no significant correlation between this measure and subjective reports of arm use (based on Pediatric Motor Activity Log – Revised (PMAL-R) scores), and a moderate correlation between accelerometry scores and a clinical measure of arm function (Pediatric Arm Function Test (PAFT)). However, Sokal’s study did not compare arm accelerometry to any concurrent objective measure of children’s arm use, as was required to validate this tool in adults. Furthermore, they did collect any normative data from typically-developing children, as is required to provide comparative standards of arm use in children.

Thus, the aim of this research project was to examine the validity of arm-worn accelerometry as a measure of arm use in children, with and without hemiparesis. The over-arching objective was to provide foundational evidence to guide future investigations of this tool’s potential as a pediatric outcome measure for children with hemiparesis. Specifically, arm accelerometry’s (concurrent) criterion validity, convergent validity and discriminative validity were examined here, in a controlled laboratory setting. Participants included children aged 3 – 14 years, both typically-developing (TD group) and with hemiparesis (CP group).

1.2 Objectives of Study

**Objective 1:** To assess the concurrent (criterion) validity of arm-worn accelerometry as a measure of bilateral arm use in children with and without hemiparesis by determining the level of agreement between accelerometry data (activity counts) and video-based observer ratings of each child’s arm movements.
**Hypothesis 1:** It was hypothesized that accelerometry would have concurrent criterion validity, as demonstrated by high accuracy (i.e. >80% agreement and Prevalence-adjusted Bias-adjusted Kappa (PABAK) scores > 0.85) in its quantification of arm usage, when compared to arm use quantities calculated based on direct observations of concurrently videotaped arm movements (considered to be the gold-standard assessment for detecting the presence of purposeful movements).

**Objective 2:** To test the convergent validity of arm-worn accelerometry as a means to measure affected arm use in children with hemiparesis by correlating arm accelerometry data (arm use ratios) with: i) scores on the Quality of Upper Extremity Skills Test (QUEST);\textsuperscript{34} and ii) scores on the PMAL-R\textsuperscript{39} (a subjective parent report measure of children’s arm use during daily activities). This objective also included the performance of additional exploratory analyses to determine if there were any correlations between accelerometry-based arm ratios and children’s Developmental Disregard Index (DDI) scores and Mirror Movement (MM) scores.

**Hypothesis 2:** It was hypothesized that accelerometry would have convergent validity, as demonstrated by moderately strong correlations (i.e. Spearman’s r >0.4 and <0.6) between accelerometry-based bilateral arm use ratios and scores on a clinical measure of arm function, the QUEST.\textsuperscript{34} Based on the QUEST’s evaluation of some constructs that are different than arm activity level (i.e. the range and quality of dissociated movements, grasp, weight-bearing, and protective extension reactions), only a moderate level of correlation was expected. Additionally, in comparing this study’s findings with that of a similar study by Sokal et al. (2015), the correlation between the accelerometry data (arm use ratios) and a subjective parent-report questionnaire targeting daily use of the affected arm (the PMAL-R)\textsuperscript{39} was hypothesized to be moderate. As well, it was hypothesized that a moderately strong negative correlation would be found between children’s accelerometry-based arm use ratios and their scores on the DDI\textsuperscript{11,57}, a composite score that estimates a child’s level of disregard of the affected arm.

**Objective 3:** To test the discriminative validity of arm-worn accelerometry as a measure of affected arm use in children with hemiparesis by: examining the differences between arm use ratios (derived from accelerometry data) between sex- and age-matched pairs of TD children and children with hemiparesis (CP group).
**Hypothesis 3**: It was hypothesized that there would be significant differences in the accelerometry-based ratio of affected to unaffected arm activity between sex- and age-matched pairs of children with hemiparesis (CP group) and without hemiparesis (TD group).

### 1.3 Potential Impact of this Study

The value of this study is that it provides a foundational level of validity testing for accelerometry as a measure of arm use in children with hemiparesis and TD children. By comparing arm accelerometer data to video-based observer ratings of arm movements in a controlled environment, the accuracy, sensitivity and specificity of arm accelerometry in measuring arm use can be evaluated. This groundwork is necessary to inform future studies which may explore arm accelerometry’s potential as a tool to discriminate between children with hemiparesis who are at different levels of arm function, and as a possible measure of their treatment intensity and treatment outcome. If arm accelerometry is successfully validated as a tool for measuring children’s arm use, it can assess a quantitative aspect of function that is targeted by treatments but not yet captured by current standardized outcome measures. In so doing, arm accelerometers may one day help to better inform best practice in paediatric rehabilitation for children with hemiparesis.
Providing further background rationale for this study, this chapter describes hemiparesis in children, outlining current physical interventions and treatment outcome measures available. The potential role for arm accelerometry as a quantitative outcome measure in children with hemiparesis is explained, with reference to elements that are missing from current outcome measures. As well, arm accelerometry’s potential value as a measure of children’s daily arm use is supported with reference to findings validating its use for this purpose in adults, post-stroke. Finally, the aims of the current study are outlined, leading into a description of this study’s potential impact.

2 Hemiparesis

2.1 Prevalence and Significance of Childhood Hemiparesis

Hemiparesis affects one third of children with cerebral palsy (CP), which is the most common paediatric physical disability, affecting more than 500,000 children and families in North America. The term cerebral palsy is used to describe a group of permanent disorders of the development of movement and posture, causing activity limitation, that are attributed to nonprogressive disturbances that occurred in the developing fetal or infant brain. At least 70-80% of children with CP present with spastic forms of the disease. Hemiparetic CP is a subtype of spastic CP that often results from a perinatal stroke and occasionally from a unilateral brain malformation. Children with spastic hemiparetic CP present primarily with asymmetric hand use secondary to weakness, decreased selective motor control and abnormal muscle tone (i.e. hypertonicity) in the affected limb(s), with occasional tremors. Although these issues are significant, most children with spastic forms of hemiparetic CP still have the capacity for some voluntary arm movements, which can be further enhanced through therapy. Only a minority (10 – 20%) of CP diagnoses are athetoid or dyskinetic types, involving involuntary movements that can be disruptive to the production of quantifiable functional arm use.

Childhood hemiparesis can also result from a post-natal stroke. Childhood stroke most often occurs during the first year of life, with a reported incidence of 1.2 – 13 cases per 100,000
children.\textsuperscript{6} This post-natal form of stroke is much rarer than CP (which often occurs due to a perinatal stroke), however it is also often under-diagnosed.\textsuperscript{2} Among children with post-natal stroke, 94\% present with hemiparesis/hemiplegia, regardless of the infarct type.\textsuperscript{6} Although children with post-natal stroke present with upper extremity deficits that are more apparent initially (compared to those with perinatal stroke), longer term functional outcomes in the upper extremity are similarly variable for both groups.\textsuperscript{7} Both of these diagnostic groups thus often receive upper extremity therapy with similar outcome goals, customized based on their individual characteristics. Thus, the quantitative measurement of targeted gains in arm use following upper extremity therapy is relevant to both children with post-natal stroke and hemiparetic CP.

Regarding the functional impact of childhood hemiparesis, this physical condition produces sensorimotor impairments that can lead to activity limitations in daily life.\textsuperscript{8,9} Children with hemiparesis often present with both a limited capacity for affected arm use, as well as levels of arm use that are even lower than those expected based on their physical capacity.\textsuperscript{10} This phenomena, termed ‘developmental disregard’, is believed to result from both operant conditioning and neuro-motor development factors. According to operant conditioning principles, the child develops learned non-use based on his or her negative experiences (i.e. pain, difficulty, failure) with previous attempts at using the affected arm. In addition, the neuro-motor explanation is that the child’s cortico-spinal projections innervating the affected arm are lacking in integrity and optimal development, due to a lack of sensory and motor stimulation through movement during critical early life periods.\textsuperscript{11}

Overall, hemiplegia has a significantly negative effect on a child’s quality of life and self-concept.\textsuperscript{12} This condition can lead to lifelong consequences, such as dependence in activities of daily living, reduced function at school, and unemployment.\textsuperscript{13,14} Specifically, in comparing children with hemiplegic CP to other forms of CP, the limited bilateral upper extremity use is the feature often associated with decreased abilities or decreased quality of performance in self-care and other activities of daily living.\textsuperscript{15,16} It has also been found that upper extremity deficits have a more adverse effect on a child’s self-esteem than lower limb deficits.\textsuperscript{17}
2.2 Therapeutic Interventions for Childhood Hemiparesis

Standard interventions for paediatric hemiparesis include Constraint-Induced Movement Therapy (CIMT) and unimanual and bimanual intensive training. CIMT involves forced restriction of the child’s unaffected arm during intensive motor practice with the hemiplegic limb, with the goal of facilitating motor learning in the affected arm. This approach has been the focus of many randomized controlled trials (RCTs) and a 2009 systematic review confirmed its short-term efficacy in children, despite concerns regarding the long-term maintenance of functional gains. However, a more recent review indicated that CIMT’s treatment effects in children are only weakly superior to those of other therapies, when the highly intense dose of treatment is controlled for. Also based on the recognition of the importance of highly intense ‘massed practice’, Virtual-Reality (VR) interventions are being utilized to target upper extremity functional gains in children with hemiparetic CP. Although the research base for VR interventions for hemiparetic CP has yet to expand, a 2010 systematic review indicated promising findings in terms of functional outcomes for this population. The rationale for this application of VR is based on evidence from adult stroke populations that therapeutic outcomes of upper extremity use are enhanced with frequent practice opportunities and engaging, motivating practice contexts.

In addition, bimanual intensive training (BIT) of the upper limbs is a well-researched form of therapy for childhood hemiparesis. Specifically, a controlled trial by Gordon et al. (2011) found that this form of training resulted in upper extremity functional gains that were equal to those achieved with CIMT, and greater than those for CIMT in terms of achievement of children’s self-identified goals. A systematic review comparing these two forms of treatment made somewhat similar conclusions, in finding that BIT and CIMT resulted in similar improvements in UE capacity and function. The differences identified were that CIMT potentially improved unilateral performance more, and BIT potentially resulted in greater bimanual performance and achievement of self-determined life goals.

Another form of therapy, Functional Electrical Stimulation (FES), has been recently investigated as part of a more comprehensive systematic review, as a means of achieving upper extremity functional outcomes for children with hemiparesis, in combination with goal-directed
FES involves the use of transcutaneous electrical stimulation, to help target specific patterns of motor activation, in the context of self-initiated voluntary movement performed by the patient. Specifically, Kapadia et al. (2013) found that a group of four post-stroke children (aged 13 years) with severe chronic hemiparesis, improved significantly on standardized measures of functional reaching and grasping, after receiving FES sessions three times weekly for 16 weeks. As well, Xu et al.’s controlled clinical trial (2012) found that the addition of electrical stimulation to the treatment protocol for children with hemiplegic CP resulted in greater gains in upper extremity function, maintained at 6-months post-intervention. More recently, Musselman et al.’s case study (2017) found both qualitative and quantitative post-treatment gains in upper extremity function for a 2-year old child, following an intense block of FES therapy. Despite these recent promising findings, further research on this form of therapy is required before more conclusive statements about its efficacy with child populations can be made.

Regardless of the intervention approach, benchmark goals of therapy are typically improved ROM, strength, motor control/dexterity, and capacity for functional arm use. However, the overarching objective is for the child to apply and incorporate his/her capacity for improved arm use into activities of daily living, within natural environments, on a long-term basis.

### 2.3 Treatment Outcome Measures for Childhood Hemiparesis

Standardized tools are available for measuring upper extremity outcomes of treatment for children with hemiparesis. As displayed in Table 1, these measurement tools can be classified based on the constructs they target (i.e. arm use capacity and function vs. arm usage during daily life), the activity context (i.e. bimanual vs. unimanual arm activity), and the type of measurement used (i.e. clinician observations, subjective reports, or objective quantification). Clinic-based tests can be sub-divided into: those measuring upper extremity function in a bilateral activity context, such as the Assisting Hand Assessment (AHA), those measuring arm function in a unimanual activity context, such as the Melbourne Assessment 2 (MA2) and the Quality of Upper Extremity Skills Test (QUEST), and those including both bilateral and unimanual activities, such as the Paediatric Arm Function Test-Revised (PAFT), as displayed in Table 1.
Specifically, the AHA evaluates the child’s ability to use the affected arm in the context of play-based bilateral activities, with evidence found for its construct validity, intra- and inter-rater reliability, and responsiveness to change in children with hemiplegic CP, between 18 months and 12 years of age.\(^3\) During the assessment, children are engaged in either semi-structured play or a theme-based board game activity (for ages > 5 years), using objects including bracelets, a crown, a medal necklace, a wind-up toy and a toy vehicle, with the objective of eliciting spontaneous use of the affected arm and hand. The AHA test administration is videotaped for offline scoring by a trained therapist. Scoring therapists rate the affected hand’s movements on an ordinal 4-point scale (with 1 indicating the child did not use the affected hand, and 4 indicating the child used the hand completely effectively). These scores are assigned for a variety of domains of object-related actions including: general use (of affected arm), arm use, grasp and release, fine motor adjustments, coordination, and pace.\(^3\)
In contrast to the AHA, the MA2 evaluates the child’s ability to use the affected arm during one-handed activities. The original MA was validated for use with children aged 5 to 15 years, and its modified version (MA2) was validated for use with children aged 2 to 4 years. A 2016 systematic review found strong evidence for MA2’s validity and reliability, however no conclusions were made regarding its responsiveness to change. The MA2 examines these elements of upper limb quality of movement: range, accuracy, dexterity, and fluency. It is also administered in the context of a table-top play session, using materials including a crayon, cube, a hairbrush, and a small (food-like) pellet. MA2’s administration is also videotaped for offline scoring by a trained therapist. Items are scored on an ordinal 3-, 4- or 5-point scale (varying between test items), and scores are calculated for each child in the domains of: reaching/to, grasping, releasing, and manipulating simple objects.

The QUEST was designed to evaluate the quality of upper extremity movements, also within a unimanual context, and it has been validated for use with children between 18 months and 8 years of age. Further, it was found to have strong test-retest reliability and moderate to strong intra- and inter-rater reliability in children with CP, aged 2.5 to 4 years. Specifically, the QUEST involves testing of these four domains of affected arm movement, resulting in a summary score for each one: Dissociated Movement, Grasp, Protective Extension, and Weight-bearing. Test activities include non-functional (e.g. assessing the child’s active ROM and ability to perform isolated hand or finger movements) and functional tasks (e.g. crayon grasp, and pellet grasp/release), as well as the testing of protective extension responses on the affected side, and the ability to co-contract statically and dynamically weight bear through the affected upper limb in a 4-point position. This test’s administration is not videotaped, and scoring is done by the administering therapist, based on observation. For each test item, a score of 1 or 2 is given, based on the quality of joint positioning or movement. Scores are summed and standardized for each domain, yielding a summary score (from 0 – 100%) for the child in each one, as well as a total summary score for all domains tested.

Another clinic-based measure, the PAFT, evaluates a child’s capacity for use of the affected arm within the context of structured unimanual and bilateral activities. Although it was developed for use with children between the ages of 2 and 12 years, its convergent validity (with arm impairment grade) was only established for children aged 2 to 6 years and its test-re-test
reliability was established for children aged 2 to 8 years.\textsuperscript{51} The PAFT administration involves engaging the child in functional play activities including grasping a puzzle knob, pulling a toy apart, pouring out of a cup, and reaching to, carrying, releasing, and throwing balls, as well as spoon use and donning a hat and boots. Test items are administered twice: the first time, with no instructions to use the affected arm/hand (to elicit spontaneous use), and the second time, with instructions (and physical cueing, if necessary) for the child to use the affected arm. The test is videotaped and scored by an independent scorer, who rates the affected arm’s movements on a 6-point scale (0 = no affected arm movement, and 5 = normal arm movement). Each child gets a total score for the: i) amount of spontaneous affected UE use; and for the ii) quality of affected UE use. As well, each child gets a score evaluating unilateral UE use, bilateral UE use, and a summary score, taking both of those into account.\textsuperscript{78}

All of the above-mentioned assessments involve child-therapist interactions in a contrived clinic setting, where the child may be aware that he/she is being evaluated. All of these assessments also rely partially on therapist interpretation for scoring. Therapist training is required for scoring the AHA and the MA2, which both involve video analysis of the child’s arm movements. Self-study is permitted for administration and scoring of the PAFT, which also involves video analysis by an assessor who is blind to the child’s treatment conditions. In contrast, there is no training protocol for the QUEST, and scoring is performed directly during testing. Regardless of the method, all four of these clinical tests incorporate a level of subjectivity (i.e. therapist interpretation) into their measurements. As well, the presence of the therapist during these assessments imparts a level of observation that may affect the child’s attempts at using the hemiparetic arm. Both of these features differentiate these assessments from accelerometry, which can potentially provide an objective quantitative measure of arm use in a natural setting. As well, all of these assessments involve eliciting the child’s best capacity for use of the affected arm, rather than the child’s actual daily arm use. However, there are consistently reported discrepancies between performance on these tests and the child’s continued use of the affected arm in daily life.\textsuperscript{38,39} These discrepancies motivated the development of another category of outcome measures, which aim to measure arm and hand usage in the child’s natural environments. Most of these measures rely on subjective parent or child report to estimate the child’s affected arm use during daily life. For example, the Pediatric Motor Activity Log Revised (PMAL-R)\textsuperscript{39} was designed to address this need, by providing parents with a questionnaire
regarding their child’s at-home arm use during a subset of typical daily activities. The PMAL-R asks parents to rate their child’s affected arm/hand use in daily life, on each 22 different items (e.g. feeding, opening a door, turning a book page). Parents rate the child’s arm use on a 6-point scale, in terms of how often it is used (HO scale) and how well it is used (HW). 79

In addition to the PMAL-R, the Children’s Hand-use Experience Questionnaire (CHEQ) 40 was designed to evaluate the child’s arm use during daily bimanual activities, based on either child or parent report. It was initially developed for children aged 6 – 18, however a MiniCHEQ was subsequently created for children < 6 years old. It is administered in the form of an online questionnaire to parents or children who are asked to rate the child’s arm/hand function during each of 27 items (e.g. pulling up a zipper jacket, handling playing cards). The test includes ratings for the child’s perceived efficacy with affected arm/hand use, the relative time it takes to complete tasks, and how much the child is bothered by the impaired hand function during each task. 40 The CHEQ’s validity and test-re-test reliability has been established for children with unilateral CP, aged 6 to 18 years. 80

Although both of the PMAL-R and CHEQ questionnaires examine a construct that is missing from clinic-based assessments, they are subject to (parent or child) recall bias and are lacking in objectivity. As well, their content is limited to a specified set of questions about a group of activities that are considered to be universal for children in the specified age ranges. Therefore, they do not account for the reality of high variation and individual differences in child routines, and they do not provide an accurate measure of exactly what each child is doing with the affected arm during his or her daily routine.

As a measure of a related construct, the Developmental Disregard Index (DDI)11,57 was designed to measure the level of discrepancy between a child’s capacity for use and his/her actual daily use of the affected arm. The DDI is a composite score that is calculated as the difference between the child’s PMAL-R score (How Often Scale) and his/her clinic-based QUEST sub-score (for Dissociated Movement). Regarding this measure, Sutcliffe et al. (2009) found that a group of five children’s DDI scores correlated with their change scores on the AHA, PMAL, and grip strength measures following a program of modified constraint therapy. 57
Arm accelerometry aims to measure the same construct of arm and hand usage in natural environments; however it is based on objective and quantitative information. Thus, it has value in its potential to provide information that is missing from both clinic-based assessments of arm function and subjective report measures of a child’s affected arm use. In particular, it can address the construct of the child’s daily arm use, by capturing information in natural environments. As well, it can provide quantitative and objective measurements that are not subject to recall bias.

Although accelerometry has not yet been validated as a measurement of children’s daily arm use, use of a similar wearable wrist monitor (measuring the frequency and range of wrist extensions) was tested in children with hemiplegic cerebral palsy. This device was able to detect wrist extensions with strong reliability (intraclass correlation coefficient, r=0.88; p<0.001) during children’s (n=15) performance of the AHA assessment. A significant correlation was found between the children’s scores on the AHA and the frequency (r=0.80; p=0.001) and range (r=0.70; p=0.008) of their wrist extensions/flexion during the test. Although this study explored a measure that was similar to accelerometry in its ability to objectively quantify arm/hand use, it did not examine the use of a device for detecting arm movements during natural play, as would be the intended application of arm accelerometry under consideration in this study.

2.4 Accelerometry

A potential tool for objectively measuring children’s arm activity in daily life is presented by accelerometry. This technology involves the use of a piezoelectric sensor that can be strapped onto each arm like a wrist-watch. Accelerometers can measure movements in up to three planes (i.e. tri-axial models), summing the activity counts to provide an overall movement vector. The information captured by these tools can be downloaded to a software program for processing, which allows for measurement of activity counts, indicating intensity and duration of arm movements over pre-specified lengths of time. Accelerometers are portable and can be used to collect movement data for several days at a time, within any setting (excluding underwater environments). Accelerometers do not provide any information regarding the quality of movement (i.e. hand dexterity) performed. However, they can potentially be used to provide information regarding the quantity of arm use, as a supplement to the information about movement quality that is captured with the AHA, MA2, QUEST or PAFT-R.
2.5 Validation of Accelerometry as a Measure of Arm Use

The initial studies that investigated accelerometry’s validity as a measurement of arm use involved adult participants with hemiparesis. Specifically, Uswatte et al. (2000) first provided evidence that an arm accelerometer (with a threshold filter) could provide an accurate measurement of real-world arm use in post-stroke adults with hemiparesis. Initially, they videotaped participants while they were wearing wrist accelerometers in a lab setting. Researchers reviewed the videos offline and indicated whether or not they observed arm movements in participants during intervals corresponding to the epoch setting of the accelerometers (i.e. set at 2-second intervals). Using these methods, they found a 93% level of agreement between the accelerometer output and their video-based observer movement scores (for dominant arm movements), first establishing arm-accelerometry’s criterion validity in post-stroke adults with hemiparesis. Since that time, arm accelerometry has been extensively studied and further validated in adults as a measure of post-treatment arm use within natural environments, as confirmed by at least two systematic reviews. 43,44 One review additionally confirmed arm accelerometry’s test-retest reliability with post-stroke adults, however conclusions about its responsiveness to change were not made.44

Subsequent studies have also found evidence for arm accelerometry’s convergent validity in adults, finding correlations between accelerometer data and clinical measures of arm use (i.e. convergent measures) to be moderately high at r = 0.6. 45 In terms of the most accurate way to calculate arm-worn accelerometer data, several adult studies have established that an arm use ratio (more affected/less affected arm use) has a higher correlation with other clinical measures of arm function, compared to the movement data obtained from only the more affected side.43 Use of this bilateral arm movement ratio is rationalized based on its ability to control for variations between participants in overall activity level, thus targeting the amount of affected arm use, relative to each person’s use of the unaffected side.

Although arm accelerometry has been established as a valid outcome measure of arm use in post-stroke adults with hemiparesis, it has not yet been validated for this purpose in children with hemiparesis. The paucity of research on using arm accelerometry as an outcome measure for children with hemiparesis is surprising, especially considering both
adults and children with hemiparesis often undergo similar forms of post-stroke upper extremity intervention, such as CIMT \(^\text{46}\) and FES.\(^\text{47}\)

Regarding accelerometry in children, research has been done to validate its use for measuring physical activity levels, with placement of devices on the lower body \(^\text{48}\) and at wrist level.\(^\text{49}\) However, only one study to-date has explored the use of accelerometers to specifically measure arm use in children with hemiparesis.\(^\text{50}\) This sole study involved 28 child participants (ages 2 – 6 years) with hemiparesis. Accelerometry data were collected from these children while they wore bilateral wrist devices at home for three days. The accelerometry data were then compared to a clinical measure of motor capacity, the Pediatric Arm Function Test (PAFT)\(^\text{51}\) and the PMAL-R,\(^\text{39}\) a measure of the child’s everyday arm use based on subjective parent report. Data were collected as a ratio of more affected arm activity to less affected or unaffected arm activity, with the rationale that this ratio was validated as an index of real-world upper extremity rehabilitation outcome in adults with stroke.\(^\text{52}\) Sokal (2015) found no significant correlation between the arm activity ratio (from accelerometer data) and subjective parental reports of the child’s arm use (based on the PMAL-R), however this measure is subject to recall bias and has no objective components. Furthermore, the PMAL-R results did not correlate strongly with the clinical measure of the children’s capacity for arm use \((r=0.43)\) in this study, further bringing into question this tool’s appropriateness as a standard of comparison for the accelerometer data. This study did find a moderate correlation between the child’s motor capacity for arm use (PAFT results) and the intensity of affected arm movements in daily life, as measured by arm accelerometers \((r=0.55)\).\(^\text{50}\)

However, in using a primary outcome measure that is based on subjective parent report (i.e. PMAL-R), this study failed to compare arm accelerometry data to any objective measure of arm use, as was necessary to validate its use in adults.\(^\text{53,54}\) Furthermore, the findings of Sokal et al. (2015) constitute only one study, and further research is required to investigate accelerometry’s potential as a tool to quantitatively measure arm use in children with hemiparesis.
2.6 The Need for an Objective Measure of Daily Arm Movement in Children

A paucity of research on arm accelerometry’s use in measuring arm activity in children exists and there is a need for an objective measurement tool to evaluate children’s arm activity levels in their home environments. Not only is this outcome (of increased arm use) one of the ultimate goals of intervention for childhood hemiparesis, but there is also a lack of correlation between the child’s performance on clinic-based tests of arm function and his/her continued use of the arm in daily life.\textsuperscript{38,39} Therefore, arm accelerometry has the potential to tap into an important component of function that is not covered by current standardized assessments. It can provide objective information regarding the amount of affected arm use, which can be assessed in combination with other measures that evaluate the quality of movement and hand/arm skills (i.e. AHA, MA2, QUEST and PAFT). Although accelerometry is limited in the breadth of information it can capture (i.e. it does not detect the quality of arm/hand movements or the child’s level of dexterity, nor does it account for the child’s use of compensatory movements and arm postures), it can supplement other measures by providing an objective recording the amount of arm activity (based on accelerations in three movement planes) in natural settings.

2.7 Validity and its Relevance to this Study

The primary objective of this research study is to examine the validity of using arm-worn accelerometry as a means to measure relative affected arm use in children. Specifically, the accelerometry output of interest involves quantifying how much the hemiparetic arm is moving compared to the unaffected arm. This research will help to identify the information captured by arm-worn accelerometers in children, and guide future investigations regarding this technology’s potential as a paediatric outcome measure for children with hemiparesis.

In evaluating any measurement tool or scale, many types of validity can be examined. In a general sense, validity refers to the measurement’s accuracy, or how well the tool measures a certain outcome of interest.\textsuperscript{55}

The major types of validity that are relevant to this study are criterion validity and construct validity. Criterion validity establishes whether a measurement tool agrees with the results of a ‘gold standard’ assessment for the same outcome. This form of validity is considered to be
Construct validity is another major type of validity. It concerns whether a measure is consistent with the theoretical construct of interest. One sub-type of construct validity is convergent validity, which addresses the correlation between a measure and another one testing the same construct. Discriminant validity is another sub-type of construct validity. It can be examined in two ways: i) by finding a low correlation between a given measure and another measure that tests a different construct; or ii) by testing whether a measure can correctly identify differences between two groups, that are known to differ on the construct being tested. Although discriminative validity is a similar term, it is specific to the measure of whether a given test can differentiate between two different groups (e.g. diagnostic groups). In this study, three types of validity were examined, as described in the aims below: concurrently-measured criterion validity, convergent validity, and discriminative validity. For the purposes of examining concurrent criterion validity here, the ‘gold standard’ measure of purposeful arm movement was video-based observation.
Chapter 3

3 Methodology

3.1 Overview

This study was a prospective cross-sectional study using quantitative methodology. With respect to the use of arm accelerometry as a measure of arm movement, this study included two phases of investigation. The first phase evaluated arm accelerometry’s concurrent criterion validity in typically-developing (TD) children. The second phase involved a preliminary analysis of arm accelerometry’s concurrent criterion, convergent and discriminative validity in a smaller sample of children with hemiparesis (CP Group). Data were collected from all child participants who attended one testing session (with parents or legal guardians present) lasting one to two hours within a single research room, equipped with a video camera, to record arm movement. For all TD participants and nine of the thirteen participants with hemiparesis, the data collection site was the Lyndhurst Centre, Toronto Rehabilitation Institute-University Health Network (TRI-UHN), in Ontario. For four of the participants with hemiparesis, the data collection site was Corbett Hall, in Edmonton, at the University of Alberta. The same procedures, room set-up, equipment, and researchers were used at both of these sites. This study was approved by the UHN Research Ethics Board (REB) and by the University of Alberta’s Research Ethics Board. Additionally, this study was approved by the Bloorview Research Institute’s Research Ethics Board, through the Toronto Academic Health Science Network (TAHSN).

3.2 Participants

Typically-developing child participants were recruited based on the following inclusion criteria: 1) aged 3-16 years, 2) had no developmental condition or other diagnosis affecting the arms. Consent was obtained from each participant’s parent or legal guardian and additionally from older children (i.e. over 8 years of age) themselves who were capable of reading the child consent form and providing written and verbal consent to the researcher. As well as parental consent, verbal assent was obtained directly from young children (i.e. ages 3 – 7, and those who could not read and fully comprehend the child consent form) using a Standard Operating Procedure for Obtaining Assent that was established according to guidelines provided by the UHN REB.
Participants with hemiparesis were recruited based on the following inclusion criteria: 1) aged 3 - 16 years, 2) had a diagnosis of hemiparesis secondary to spastic cerebral palsy (CP) or childhood stroke, as confirmed by a medical doctor familiar with the child, 3) had the ability to sit unsupported for at least 1 minute, 4) classified on the Manual Ability Classification System (MACS) at a level I, II, III or IV, as assessed via phone interview with a parent, conducted by this researcher (J. Dawe), a registered Occupational Therapist (OT), during study screening. This last criterion was included to ensure that participants had at least a minimal level of ability to perform functional play tasks involving at least partial use of their impaired arm. The MACS describes how children with CP use their hands for object handling during daily activities. It has been validated for use with children between ages 4 and 18 years, and is based on subjective descriptions from parents and teachers who interact with the child on a frequent basis. MACS classification levels range from I (handles objects easily and successfully) to V (does not handle objects and has severely limited ability to perform even simple actions).

Exclusion criteria for participants with hemiparesis were as follows: 1) medical co-morbidities affecting motor development or sensory function of upper extremities, such as peripheral neuropathy, 2) uncontrollable seizures, 3) any significant upper extremity contractures of the shoulder, elbow, wrist, or hand, as reported by parent/guardian during the screening interview (i.e. based on parent responses indicating that the child had limited movement in any of these joints), 4) a non-spastic CP diagnosis (dyskinetic or ataxic). In rationalizing the exclusion criteria, the purpose of this study was to measure voluntary movement of the affected arm. Hence, the presence of any impairment or condition involving involuntary arm movements would have precluded measurement of this study’s variable of interest. However, participants were not necessarily excluded if they had another developmental disability. All participants with hemiparesis were recruited from a Canadian Cerebral Palsy Registry. Specifically, nine participants were recruited in Toronto, Ontario from the Hemi-NET Database (based on information provided by Dr. Darcy Fehlings and Ms. Lauren Switzer of the Bloorview Research Institute), which is a registry including children with a diagnosis of hemiplegic CP. As well, four participants were recruited, by co-investigator Dr. Jaynie Yang, from the Canadian CP Registry in Alberta, Ontario.
Regarding the total sample size, 26 typically-developing children were recruited (in Toronto, Ontario) and thirteen children with hemiparesis were recruited and tested (in Toronto, Ontario and Edmonton, Alberta). The size of the typically-developing group (n = 26) meets the recommended sample size (n = 19) that would be required to have the power to demonstrate a correlation of r = 0.6 or higher (at an alpha level of 0.05 and beta level of 0.20)\textsuperscript{59} between this study’s primary variables of interest. This estimate is based on previously published results from adult studies that found correlation coefficients between wrist-worn accelerometry data and other measures of arm use ranging from 0.60 \textsuperscript{53,54} to 0.94.\textsuperscript{53} Based on these published findings for adults, it was expected that at least a moderately high correlation (i.e. r = 0.6 or higher) between accelerometry data and videotaped observer ratings of children’s arm movements would be found. As well, this study’s control group sample size (n = 26) nearly approaches the guidelines recommended for a sample size rated as at least Fair (30 < n < 49) by Consensus-based Standards for the Selection of Health Status Measurement Instruments (COSMIN), in reference to the testing of hypotheses and criterion validity.\textsuperscript{60} Regarding the group of participants with hemiparesis (n = 13) and the study’s primary objective, the size of this group does not meet the recommended sample size (n = 19) required, assuming these same parameters (i.e. r = 0.6, alpha = 0.05, beta = 0.20). However, in the case of a higher expected correlation of r = 0.8, the recommended sample size would be n = 12, at which level the size of the group with hemiparesis (CP group) examined here would meet the recommended sample size in addressing Objective 1.\textsuperscript{59}

### 3.3 Recruitment of Participants

Typically-developing children were recruited through word-of-mouth and emails sent to “all users” at the Lyndhurst Centre, TRI-UHN, and the Neural Engineering and Therapeutics team in Toronto, Ontario. In addition, hard copies of recruitment flyers were posted in the main areas of TRI-UHN’s Lyndhurst Centre.

Several methods were utilized to recruit participants with hemiparesis. An electronic notice of the study was posted on the Bloorview Research Institute’s study recruitment web page and on the Ontario Cerebral Palsy Foundation’s web page. Recruitment flyers were also emailed to several clinics in the Greater Toronto Area that provide treatment to children with posture and movement disorders. As well, a request was made for permission to contact families registered on the Hemi-Net database, whose children met this study’s inclusion criteria. This database includes families
with children with hemiparesis who have consented to be contacted about research studies. Contact information for these families was provided by Dr. Darcy Fehlings’ laboratory at the Bloorview Research Institute. Subsequently, a recruitment letter was mailed to families, who were provided with the opportunity to opt-out of this opportunity. For families who did not opt-out, a researcher (J. Dawe) followed up with a phone call after two weeks, to request their participation in the study. Paper versions of study flyers were also distributed to clinicians at the Holland Bloorview Kids Rehabilitation Hospital, to be shared with potentially eligible families.

Finally, participants with hemiparesis in Alberta, Ontario were recruited by Dr. Jaynie Yang of the University of Alberta. Dr. Yang provided J. Dawe with the contact information of families who were eligible and interested in participating in the study. J. Dawe followed up with a telephone screening to verify that they met the inclusion criteria.

3.4 Study Procedures

For TD participants, study procedures took place in this order: 1) collection of demographic information from a parent/legal guardian (see Appendix A), 2) video-taped, semi-structured play while the child wore bilateral wrist accelerometers. Intermittent periods of arm stillness and shaking were performed in between play tasks. During arm stillness, the child was instructed to rest hands on the table top (i.e. “stay still like a statue”). During arm shaking, the child was cued to vigorously shake rattles in the air. Refer to Figure 1 for an overview of the total play time:

![Figure 1. Structure of Play Time](image)

*These intervals provided opportunities for accelerometry-based measurement of non-movement and highly intense arm movement.*
For participants with hemiparesis, the study involved the following procedures in this order: 1) collection of demographic and medical information about the child from a parent/legal guardian, 2) video-taped, semi-structured play (see Figure 1) while the child wore bilateral wrist accelerometers, 3) clinical assessment (i.e. Quality of Upper Extremity Skills Test (QUEST), fine motor screening, and Modified Tardieu Scale, examining spasticity at the elbow and wrist) of the child by Jaclyn Dawe, a registered occupational therapist (OT), and 4) parent questionnaire (Paediatric Motor Activity Log – Revised (PMAL-R)), also administered by the OT. All participants were permitted to take rest breaks, as needed.

Detailing of the study procedures listed above:

For all participants:

1) Collection of demographic and medical information: The parent or legal guardian was asked the child’s month and year of birth, sex, diagnosis (if applicable), and age of diagnosis (if stroke was acquired post-natally). Older children (i.e. ages 7 and up) were given the option to participate in this discussion as well (see Appendix A for the data collection sheet).

2) Videotaped, Semi-Structured Play in Lab: This procedure lasted about 20 minutes. During this time, each child participant wore two wireless, tri-axial accelerometers (Actigraph wGT3X-BT), with one on each wrist. Accelerometers were secured to the wrists with Velcro straps. Each child was asked to engage in play activities while sitting at a table with a researcher (J. Dawe). This period was video-taped using a tripod camera (JVC Hard Drive Camcorder, Everio GZ-MG330RU). The video camera was positioned directly in front (at an approximately 3-foot horizontal distance) of each participant, approximately one foot above the participant’s arm level on the table. An online digital clock was displayed on a laptop that was visible in the video frame. As the accelerometry data were time-stamped, the digital clock enabled precise alignment of video and accelerometry data. The tabletop play activities were semi-structured as follows: i) Initial five-minuteplay
period: activities with bimanual completion opportunities were presented to the child (i.e. beading, foosball, card shuffling, making shapes with playdough, etc.); ii) Second five-minute play period: unimanual opportunities for play activities were presented to the child (i.e. drawing, colouring, completing a block puzzle, building a tower, etc.); iii) Third five-minute play period: the child selected his/her favorite activity option within the clinical testing room, regardless of the unimanual or bimanual components required (i.e. ‘free play’). Due to the age range of child participants in this study (i.e. 3 – 14 years), different developmental-appropriate toy options were provided during each play period, based on the age and level of function of each participant. For example, some older participants chose to bring in an iPad or tabletop video game that they liked to play at home, to use during this experiment. In addition, before and after each 5-minute play period, the child was instructed to engage in 90 seconds of alternating intervals of arm stillness and arm shaking (i.e. highly intense arm movements), in the following order: i) 30 seconds of stillness, ii) 30 seconds of arm shaking, iii) 30 seconds of stillness. The purpose of these stillness/shaking intervals was to ensure that the 20-minute period of observation included enough intervals of both non-movement and high intensity arm movement conditions so that the evaluation of specificity and sensitivity of the accelerometers’ detection of arm movements would not be biased (i.e. by a lack of prevalence of instances of movement or non-movement).

3) For child participants with hemiparesis, clinical measures were also collected by J. Dawe, an OT, during the testing session in the following order:

a. The QUEST has been validated for use with children with CP between ages 18 months and eight years.\(^{34}\) It measures the quality of movements of both upper extremities. It is administered and scored by a therapist based on observations during a series of structured play activities, over a duration of 30 to 40 minutes. It evaluates movement of both arms in the domains of: dissociated movement, grasp, weight bearing and protective extension. The QUEST was administered and manually scored (during the study visit) by J. Dawe, an OT trained in its procedures. Each child’s QUEST scores were collected as outcome variables for the evaluation of convergent validity of accelerometry-based arm use ratios. For both the Grasp scale and the Dissociated Movement scales of the QUEST, bilateral
scores and unilateral (hemiparetic arm) scores were calculated for each participant with hemiparesis.

b. Fine motor screening for mirror movements: This observational screening is part of the Paediatric Stroke Outcome Measure-Neuro Exam, which was also utilized in the Hemi-NET’s Neurodevelopmental Platform – Physician (PSOM) Data Collection Sheet from Holland Bloorview Kids Rehabilitation Hospital. It involves observations of the child’s performance of three different fine motor actions (i.e. rapid index finger taps, sequential finger opposition) with each hand, separately. During this time, the assessing therapist observes the opposite hand for the presence of movement, ranging from ‘none’ to ‘movements equal to that expected for the intended hand’. Based on this information, an ordinal score (from 0 – 4) is derived for each activity, indicating the degree of presence of mirror movements in each arm. This information was collected as a descriptive measure of each participant with hemiparesis. Although mirror movement measurement was not part of this study’s main objectives, it was examined here (via exploratory analysis) as a possible correlate of accelerometry-based arm use ratios in child participants with CP.

c. The Modified Tardieu Scale (MTS) is a measure of spasticity that considers resistance to passive stretch of a muscle at slow and fast speeds. It has been used in children with cerebral palsy as young as three years of age and has adequate inter-rater reliability in this population. This test was administered by passively moving the child’s affected elbow and wrist through the available range of motion (ROM) at both fast and slow speeds. A goniometer was used to measure the angle of ‘catch’ during fast passive movements as well as the full passive ROM. Its administration takes approximately two minutes. This measure of spasticity was used to provide one measure (i.e. an upper limb spasticity rating of Grade 0 – 5, based on the level of catch angle or quality of muscle response to fast stretch, for both the affected elbow and wrist). Specifically, the following scale was used to assign Grades on the MTS scale: Grade 0, indicating no resistance to passive fast stretch; Grade 1, indicating slight resistance, with no clear catch angle; Grade 2, indicating a clear catch at a precise angle, followed directly by release; Grade 3, indicating fatiguing clonus (< 10 seconds) at a precise angle; Grade 4, indicating indefatigable clonus at a precise angle; Grade 5, indicating that a joint is immobile to
passive fast stretch. This measure was used here to provide descriptive data for each child participant with hemiparesis.

4) The Paediatric Motor Activity Log - Revised (PMAL-R) was administered to parents of child participants with hemiparesis. It is a parent-report measure of a child’s use of his/her affected arm in daily activities, such as turning a doorknob and picking up and holding objects.\textsuperscript{39} Parents are asked to rate how often and how well the activities are performed by their child with the affected arm. It is primarily used for children aged under nine years; however, it was administered to parents of children aged of 4 - 14 years in this study. For parents of children older than nine years, they did not indicate that any of the test items were activities that their children did not perform. The PMAL-R has been reported to have the following verified psychometric properties: unidimensionality, construct validity and good test-retest reliability.\textsuperscript{39,63} The interview takes approximately 15 minutes to administer. It was also reported by Uswatte et al. (2012) to have moderate convergent validity with a clinic-based measure of arm function, the Paediatric Arm Function Test (PAFT). However, it is noteworthy that Uswatte et al. also found that although change scores on this measure (i.e. post vs. pre-treatment) were consistently reliable across different participants, scores based on a one-time administration of this measure (as utilized in this study) were not reliable or valid between different participants as a global measure of arm function. This was explained by Uswatte et al. as potentially due to differences between parents in their frames of reference for estimating amounts of arm use in their children. Despite this threat to the validity of the PMAL-R as a one-time measure of arm use, it was selected for use in this study so that results could be compared to those obtained by other published studies, such as Sokal et al. (2015), which used the PMAL-R directly, and Uswatte et al. (2000), which used the adult version (MAL) of this questionnaire when testing the convergent validity of arm-accelerometry in post-stroke adults.
3.5 Data Analysis

3.5.1 Analysis of Videotape Data

The video footage was reviewed offline by one of three researcher Observers. Observer 1 was a research student (J. Dawe), Observer 2 was a research assistant (N. Unic) and Observer 3 was a research staff member (V. Zivanovic). The researchers analyzed each two-second epoch, scoring it for the presence or absence of arm movements (i.e. assigning a score of zero (absence) or one (presence) to each two-second epoch). As well, researchers scored each two-second epoch for the relative size of arm movements observed (i.e. rating movements on an ordinal scale as either: very small, small, medium or large). The following guidelines were used by observers in determining the size to assign each observable upper extremity movement during each 2-second epoch: small movements = 1/3 or less of the child’s ROM (at a given UE joint) used; medium movements = 1/3 to 2/3 of the child’s ROM used; large movements = >2/3 of the child’s ROM used, or less than 2/3 ROM, but with fast-paced movements (e.g. vigorous shaking at elbow and wrist); very small movements = movement observable on-screen (i.e. arm not still), but smaller than the majority of those classified as small. Since all child participants had full ROM at all upper extremity joints (based on this study’s inclusion/exclusion criteria), observers were able to estimate their relative UE ROM used by viewing the video data alone.

Based on observer scoring, a total score of observed arm movement was produced for each arm, by summing the number two-second epochs that included any arm movement. This observer-based score of movement for each arm was divided by the total number of epochs observed, to obtain a score indicating the proportion of epochs that included arm movement for each arm.

The video-based observations of arm movements for the majority of participants (25/26 typically-developing and 11/13 participants with hemiparesis) were completed by one researcher (J.Dawe, Observer 1). The remaining video-based observations that were used for criterion validity testing (i.e. addressing Objective 1) were completed by a research staff (Observer 3).

3.5.2 Inter-Rater Reliability Testing

For the purposes of inter-reliability testing of video-based observations only, 7 – 10-minute samples of video-based observations were completed by three independent researcher Observers:
Observer 1 (J.Dawe), Observer 2 (research student, N. Unic), and Observer 3 (research staff, V. Zivanovic). Video data used for reliability testing included footage from both typically-developing participants and participants with hemiparesis. For this inter-rater reliability testing, each sample of video-based data was independently observed by a pair of researchers (i.e. either Observer 1 and Observer 2 or Observer 1 and Observer 3). To calculate agreement between Observer 1 (J.Dawe) and Observer 2 (a research student), eight sets of data, from four different participants (typically-developing), were compared. To calculate agreement between Observer 1 (J.Dawe) and Observer 3 (research staff), four sets of data, from two different participants (with hemiparesis) were compared. Each pair’s ratings of arm movements were compared to evaluate inter-rater reliability using the prevalence-adjusted bias-adjusted Kappa (PABAK) statistic. The PABAK was chosen as the test statistic, as it adjusts for any differences in the prevalence of a given condition (e.g. an arm movement vs. no arm movement condition), which may have a biasing effect on the Kappa statistic. The PABAK also adjusts for any systematic biases in observers’ ratings. In interpreting PABAK scores, a reference scale was used (see Table 2), as reported in Knights et al. (2013).

Table 2. Interpretation of PABAK Scores, as reported in Knights et al., (2013)

<table>
<thead>
<tr>
<th>Interpretation</th>
<th>PABAK Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fair</td>
<td>0.21 – 0.40</td>
</tr>
<tr>
<td>Moderate</td>
<td>0.41 – 0.60</td>
</tr>
<tr>
<td>Good</td>
<td>0.61 – 0.80</td>
</tr>
<tr>
<td>Excellent</td>
<td>0.81 – 1.00</td>
</tr>
</tbody>
</table>

PABAK (Nominal) scores were determined based on individual observations of the presence or absence of movement in dominant arms and non-dominant or hemiplegic arms. PABAK (Nominal) scores were calculated separately for each pair of observers (i.e. Observers 1-2 and Observers 1-3).

PABAK Ordinal Scores were calculated to determine the level of agreement between pairs of researchers regarding the size of observed arm movements during each 2-second epoch of video.
Observers rated observed arm movement(s) on an ordinal scale as either: very small, small, medium, or large, during each 2-second frame of video. When deciding on a rating level, observers took into account the relative ROM used (i.e. full vs. partial), joint involvement (i.e. elbow only or simultaneous elbow and shoulder movement), and the speed and frequency of movement(s) during each 2-second interval. The PABAK Ordinal Score takes into account the proportion of times that observers exactly agree on the size of a movement. As well, it provides a 2/3 weighting for the proportion of times that observers nearly agree on the size of a movement (e.g. if Observer 1’s rating = small, and Observer 2’s rating = very small). PABAK Ordinal Scores were determined individually for observations of movement in dominant arms and non-dominant or hemiplegic arms. PABAK Ordinal Scores were calculated separately for each pair of observers (i.e. Observers 1-2 and Observers 1-3).

3.5.3 Arm Accelerometry Data Processing

Accelerometry-based data (Actigraph wGT3X-BT) were uploaded from the accelerometer monitors and processed using ActiLife 6 software (Actigraph). According to Actigraph’s specifications, the devices were set to sample accelerations at 10 Hz and summed over a specified epoch of two seconds, for each device (i.e., one worn on the wrist on the affected or non-dominant side, and one worn on the wrist of the non-affected side or dominant arm). This sum of accelerations over the user-specified epoch (i.e. 2-seconds in this study) has been referred to as a.raw count, providing a rough index of the amount of movement in the object to which the accelerometer is attached. Raw counts have been used as a general term to describe accelerations measured by uniaxial, biaxial, and triaxial accelerometers. Specific to triaxial accelerometers, the Vector Magnitude (VM) is also a quantity estimating the total amount of acceleration during the length of each epoch, combined across the 3 (x,y,z) planes of movement being measured. Actigraph 6’s software (through its proprietary processes) removes accelerations due to gravity and converts accelerations into activity counts (0.001664g/count). The accelerations during each epoch are combined across the 3 measured axes into a single Vector Magnitude (VM) value, based on the equation \( \sqrt{x^2 + y^2 + z^2} \). Therefore, the VM value, when using a 2-second epoch, represents the number of activity counts per 2-second epoch (i.e. counts/epoch).
3.5.4 Receiver Operating Characteristic (ROC) Analysis

To determine the optimal threshold setting for the accelerometer’s detection of arm movement (i.e. to identify the presence or absence of movement for each two-second epoch), ROC analysis (using SPSS, Version 24) was performed. This analysis was performed for all observer (i.e. video) and accelerometer ratings of the presence/absence of arm movement from each participant. ROC analysis enabled investigation of how the sensitivity and specificity of accelerometer measurements varied at different threshold settings (i.e. at differing values of the VM). When analyzing accelerometry-derived VMs, the optimal threshold was deemed to be the value of VM (counts/2-second epoch) at which the sensitivity and specificity of this tool’s measurements were maximized with equal weighting, based on the calculation of Youden’s Index, \( J (J = \text{sensitivity} + \text{specificity} - 1) \). The optimal threshold (the threshold value at which Youden’s Index was maximal) was calculated for data from each arm of each participant. By analyzing the distribution of all optimal threshold values in each group of participants, a median optimal threshold value was determined for the typically-developing (TD) group and for the group with hemiplegic cerebral palsy (CP). Next, a determination regarding the optimal threshold to use for future analyses (i.e. to address Objectives 1, 2 and 3) for each group was made by comparing (using a 1-way ANOVA) the mean Youden’s Index value at three different thresholds: the default threshold of zero, the median optimal threshold for the group, and a threshold of 2.0 counts/epoch. This threshold of 2.0 was included in the analysis, based on findings from previously published literature. Specifically, a few previous studies have indicated that a threshold setting of 2.0 was optimal in measuring post-stroke adults’ arm use. Uswatte et al. (2000) reported determining this by varying thresholds when filtering accelerometry data from their adult participants, and selecting a threshold value that maximized the percent agreement between accelerometer and observer arm movement (duration) scores. This optimal threshold setting of 2.0 was based on Uswatte et al.’s use of uniaxial accelerometers from Computer Science and Applications, Inc. (model 7164). Sokal et al. (2015) subsequently applied use of this threshold of 2.0 for analyzing arm accelerometry output from biaxial Actigraph accelerometers. More recently, Waddell et al. (2017) using Actigraph 6’s proprietary software for triaxial accelerometers, also applied use of a threshold of 2.0, when interpreting their accelerometry-based data in measuring the upper limb performance of adults with hemiparesis.
3.5.5 Accelerometry-based Arm Movement Ratios

Arm use ratios were calculated for each participant (in both TD and CP groups), based on the proportion of all the 2-second epochs (during the ~ 20-minute period of videotaped tabletop activity) that included accelerometry-based detection of arm movement (i.e. VM > optimal threshold setting), compared between the non-dominant (or hemiparetic) and dominant arm, for each participant. Specifically, arm use ratios quantified the duration of movement in the non-dominant arm (or hemiparetic arm) relative to the dominant arm, over the recording period. Arm ratios were calculated by summing the number of epochs with above-threshold VM values for the non-dominant (or hemiparetic) arm, and then dividing that number by the number of epochs with above-threshold VM values for the dominant arm, for each participant. This calculation was based on an equal number of recording epochs for each arm, since the accelerometers were worn and recording simultaneously. The resultant ratio was converted to a percentage value (out of 100). Thus, a ratio of 1:1 was converted to 100% (indicating equal duration of movement in both arms) and a ratio of 1:2 was converted to 50% (indicating that the non-dominant or hemiparetic arm was moving for half the duration of the dominant arm’s movements).

For the TD group of participants, arm ratios were compared between different age groups with a 1-way ANOVA. For this analysis, participants were categorized into one of the following age groups: 3-year olds; 4-5 year olds; 6-7 year olds; 8-10 year olds. The reason for choosing uneven age ranges for the youngest (1-year range) and oldest (3-year range) categories was because of the greater rate of development that occurs earlier in life. Specifically, there is a greater expected developmental difference between 3- and 4-year olds, than between 9- and 10-year olds. Therefore, it was considered appropriate to have a smaller age range for the younger age group, and a larger age range for the oldest age group. As well, this categorization decision maximized the equality between the number of participants in each group.

3.5.6 Evaluation of Criterion Validity

To address Objective 1 of this study, the ability of the accelerometers to detect the presence/absence of movement was compared to the gold-standard measure (i.e. video observation). The absolute agreement was calculated for each arm of each child in both the typically-developing group and the group with hemiparesis (i.e. CP group). Agreement was
calculated as the proportion of epochs in which there was agreement between the video and accelerometry data regarding the presence or absence (i.e. a VM score below the optimal threshold for movement detection, and an observer score of 0) of arm movement. The mean agreement was calculated for the dominant and non-dominant arms of typically-developing children, and the dominant and hemiparetic arms of the children with hemiparesis. In addition, the statistical significance of agreement between observer and accelerometry was also determined by calculating the PABAK. Agreement levels (% agreement and PABAK coefficients) were initially calculated for each arm, and compared across the three accelerometry thresholds investigated: the (default) zero threshold, the median optimal threshold for that participant group (TD or CP), and a threshold of 2.0. As well, a 2-way ANOVA was used to analyze the difference between dominant and non-dominant arm movement agreement levels, at each threshold, for each participant group.

3.5.7 Disregard Index Calculation for Participants with Hemiparesis

A Disregard Index for each child with hemiparesis, as utilized in previous research involving children with hemiplegic cerebral palsy, was calculated. The Disregard Index is a quantitative estimate of the extent to which the child disregards his/her affected arm during daily activities, based on parent report. It is the difference between the child’s score on the QUEST (which assesses his/her ability to move and use the affected arm under ideal conditions) and the child’s score on the PMAL-R (which estimates the amount of affected arm use, based on parent report, on a typical basis in the child’s daily life). The Disregard Index = QUEST score for Dissociated Movement (in % form) – [PMAL-R score (How Often Scale)/5 x 100%]. The Disregard Index value is a number, indicating the presence of affected arm disregard (i.e. values greater than zero) or the absence of arm disregard (i.e. values less than or equal to zero). As part of an additional exploratory analysis, DDI scores for each child participant with hemiparesis were examined to see if they were correlated with accelerometry-based arm use ratios. Although this analysis was not one of this study’s main objectives, it was hypothesized that children’s DDI scores (indicating the level of disregard of the affected arm) would be negatively correlated with their accelerometry-based arm use ratios.
3.5.8 Evaluation of Convergent Validity

To evaluate the convergent validity of accelerometry to measure arm movement in children (i.e. Objective 2), arm use ratios were calculated for each participant in the CP group (see Section 2.5.5). Use of this ratio summary variable has been validated as a measure of relative arm use in adults with post-stroke hemiplegia.52

To evaluate the convergent validity of arm-worn accelerometry as a measure of relative hemiparetic arm movement, the arm-movement ratio was correlated with the sub-scores on the QUEST (i.e. Dissociated Movement scale – with an analysis of both bilateral and unilateral scores) and the revised P-MAL (i.e. How Often scale). The Dissociated Movement scale of the QUEST was selected as a potential correlate of arm movement ratios because this scale provides a measure of both active ROM at upper extremity joints and selective motor control (e.g. the ability to maintain a neutral wrist and extend fingers with the upper arm in different positions). It was hypothesized that greater active ROM and selective motor control in the hemiparetic arm would be associated with increased hemiparetic (relative to non-hemiparetic) arm movement during tabletop activities. In addition, the PMAL-R was selected as a potential correlate of arm movement ratios, because this measure targets actual hemiparetic arm use during daily life. As well, Sokal et al.'s study (2015) examined the correlation between children’s arm ratios and their PMAL-R scores. Therefore, use of this clinical measure allowed for a comparison of the current results to those found by Sokal et al.

As an additional exploratory analysis, arm use ratios were tested for correlation with QUEST Grasp scale scores (bilateral and unilateral), Mirror Movement scores (bilateral), and Developmental Disregard Index (DDI) scores. First, the data were tested for normality of distribution using the Shapiro-Wilk test (recommended for smaller sample sizes of n < 50).69 Results of this test indicated that the distribution of data points for the QUEST scores, arm use ratios, PMAL-R scores, and DDI scores were not significantly different than normal distributions. As well, levels of skewness and kurtosis in the distributions of these variables were calculated and found not to exceed the assumptions of normality (i.e. ratios of skewness and kurtosis to the standard errors of these values were less than 2). Since the independent variable in the correlational analyses (i.e. arm use ratio) was a continuous ratio variable and the primary clinical test variables were ordinal (i.e. QUEST and PMAL-R scores), it was determined that the non-
parametric Spearman’s Rho (rₚ) would be an appropriate test to use. For the additional exploratory correlational analysis involving the children’s Mirror Movements scores as a dependent variable, Spearman’s Rho was also used, due to the non-normal distribution of this sample’s Mirror Movement scores, based on the Shapiro-Wilk test.

Regarding the independent variable (arm use ratios), correlational analysis was performed twice, to include arm use ratios calculated at two different threshold settings (i.e. i) arm use ratios at the optimal threshold (2.0) for detection of all movement; and ii) arm use ratios based on a higher threshold of movement detection that eliminated the inclusion of very small movements). This decision to exclude very small movements was made based on the observation that very small movements were primarily non-functional and involuntary (i.e. during intervals when both typically-developing children and those with hemiparesis were instructed to remain still, they would sometimes produce observable very small observable arm movements). These very small movements did not appear to differ in frequency between children with and without hemiparesis. However, when child participants produced larger movements (i.e. using greater ROM and at faster speeds), differences between those with and without hemiparesis were markedly observable. Therefore, it was expected that differences in their arm use ratios would be revealed by excluding the very small movements (at a lower VM based on accelerometry output). Similarly, Uswatte et al (2000) also previously found that by excluding intervals with very small arm movements, they could maximize the accelerometer and observer agreement when analyzing arm movements of post-stroke adults. In rationalizing their decision to use this higher threshold that excluded very small movements, they further reported finding that 92% of the intervals in which there was accelerometer-observer disagreement contained either small or non-functional arm activity.

3.5.9 Evaluation of Discriminative Validity

To evaluate accelerometry’s discriminative validity, a sub-set of the total sample data was used. Typically-developing children and the group with hemiparesis were directly matched for age and sex, and n = 9 matches were used for this analysis. The distribution of arm use ratio data did not violate assumptions of normality (i.e. normal distribution of data, homogeneity in variances at different levels of the dependent variable), and a matched pairs t-test was used to compare the mean ratios of non-dominant to dominant arm use between the children with hemiparesis (CP group) and without hemiparesis (TD group). Despite the small sample size (n = 9 per group), this
test was chosen based on the continuous nature of the outcome variable (arm ratios) and the normal distribution of data. This analysis was performed twice: for arm use ratios calculated at two different threshold settings (i.e. i) arm use ratios at an optimal threshold of 2.0; and ii) arm use ratios calculated based on a higher threshold of movement detection that eliminated the inclusion of very small movements). An alpha level of 0.05 was used for all statistical tests in this thesis, which were completed with SPSS, Version 24.
Chapter 4

4 Results

4.1 Participant Demographics

4.1.1 Typically-developing (TD) Group

The TD group consisted of 26 participants, ranging in age from 3 – 10 years. The age, sex, and hand dominance of these participants are summarized in Table 3. This sample included 12 males and 14 females. Within this sample, 21 participants were right-hand dominant (i.e. approximately 81%) and 5 participants were left-hand dominant (i.e. approximately 19%).

Table 3. Demographic Characteristics of Typically-Developing (TD) Group

<table>
<thead>
<tr>
<th>Age</th>
<th>3 years</th>
<th>4-5 years</th>
<th>6-7 years</th>
<th>8-10 years</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of Participants</td>
<td>5</td>
<td>9</td>
<td>7</td>
<td>5</td>
</tr>
<tr>
<td>Sex:</td>
<td>Male:</td>
<td>4 (3R, 1L)</td>
<td>2 (R)</td>
<td>3 (1R, 2L)</td>
</tr>
<tr>
<td></td>
<td>Female:</td>
<td>1 (R)</td>
<td>7 (R)</td>
<td>4 (3R, 1L)</td>
</tr>
</tbody>
</table>

Abbreviations: R, Right-hand dominant; L, Left-hand dominant

4.1.2 Participants with Hemiparesis – CP Group

The group of participants with hemiparesis (CP Group) consisted of 13 individuals, ranging in age from 4 – 14 years. The age, sex, diagnosis, and hemiparetic side of these participants are summarized in Table 4. This sample group included five males and eight females. All of these participants, who were enrolled in a Canadian CP registry, had received a diagnosis of CP, as confirmed by parent report. In addition, two of these participants (one male, age 13, and one female, age 9) had a concurrent diagnosis of Autism Spectrum Disorder (ASD). Nine of the 13 participants with hemiparesis were matched in age and sex to participants within the typically-developing (TD) group.
Regarding the two participants with an ASD diagnosis, these children were not initially excluded because the presence of involuntary arm movements was not indicated by their parents during the study screening stage. However, observations during data collection indicated that they both presented with motor overflow and involuntary movements in both arms. The presence of such involuntary arm movements was one of the exclusion criteria. Therefore, data from these participants were excluded from the final analysis addressing Objective 2 of this study. The other reason for excluding their data was because Objective 2 involved analyzing participants’ QUEST and PMAL-R scores. These clinical measures have not been validated for use with children with an ASD diagnosis. Data from participants with hemiparesis and ASD were still included in the analyses addressing Objective 1 and 3, as addressing these objectives did not involve using clinical measures designed for children with only hemiparesis. As well, Objective 1 did not involve an analysis of participants’ accelerometry-based arm use ratios, which would have been affected by the presence of involuntary arm movements.

Table 4. Demographic Characteristics of Children with Hemiparesis (CP Group)

<table>
<thead>
<tr>
<th>Age Range</th>
<th>4-5 years</th>
<th>6-7 years</th>
<th>8-10 years</th>
<th>11-14 years</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of Participants</td>
<td>3</td>
<td>3</td>
<td>4</td>
<td>3</td>
</tr>
<tr>
<td>Sex: Male:</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>Female:</td>
<td>3</td>
<td>3</td>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td>Hemiparetic Side</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Right:</td>
<td>2</td>
<td>1</td>
<td>3</td>
<td>2</td>
</tr>
<tr>
<td>Left:</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td>1</td>
</tr>
</tbody>
</table>
4.2 Inter-Rater Reliability of Video Observations

4.2.1 PABAK (Nominal Scores)

PABAK coefficients of agreement regarding the presence/absence of dominant arm movement were 0.86 (Observer 1 and Observer 2) and 0.95 (Observer 1 and Observer 3). For non-dominant arm movements, PABAK was 0.79 (Observer 1 and Observer 2). For hemiparetic arm movements, PABAK was 0.84 (Observer 1 and Observer 3). For both observer pairs, levels of agreement were higher regarding dominant arm movements compared to non-dominant/hemiparetic arm movements. Based on the rating scale used by Knights et al. (2013) in Table 2, all PABAK (Nominal) coefficients were within the Good – Excellent ranges. This result indicates that video-based observations of arm movement were acceptably equal between different observers, and that similar results would be expected if another observer were to perform these video-based observations of arm movement.

4.2.2 PABAK (Ordinal Scores)

PABAK (Ordinal) coefficients of agreement regarding the size of dominant arm movements were 0.52 (Observer 1 and Observer 3) and 0.71 (Observer 1 and Observer 2). For non-dominant and hemiparetic arm movements, PABAK (Ordinal) coefficients were 0.65 (Observer 1 and Observer 3) and 0.72 (Observer 1 and Observer 2). For both observer pairs, levels of agreement were slightly lower regarding the size of dominant arm movements (0.521 – 0.71) compared to ratings of the size of non-dominant or hemiparetic movements (0.649 – 0.717). Based on the rating scale used by Knights et al. (2013) in Table 2, the majority of PABAK (Ordinal) coefficients were within the Good range, with the exception of Observer 1-3 agreement regarding the size of dominant arm movements (0.521), which was in the Moderate range.

4.3 Criterion Validity of Accelerometry - Typically-developing (TD) Group

In addressing Objective 1, accelerometry’s accuracy in identifying instances of observed arm movement (i.e. true positives) and no observed arm movement (i.e. true negatives) was determined, by calculating sensitivity and specificity values based on data from each arm for each participant. Initial sensitivity and specificity values were determined at a default VM threshold setting of zero. At a zero threshold setting, mean sensitivity was higher (90.59 +/- 5.62% for
dominant arms; 83.61 +/- 16.8% for non-dominant arms) than mean specificity (74.1 +/- 17.32% for dominant arms; 74.22% +/- 14.12% for non-dominant arms), for all 26 TD participants.

4.3.1 ROC Analysis: Determining an Optimal Threshold Setting

Evaluating sensitivity and specificity at different threshold settings, ROC analysis and Youden’s Index were used to determine optimal threshold setting values for data from each arm for each participant. Based on this analysis, it was found that optimal thresholds were most frequently between VM values of 0 and 5, as shown in the Frequency Distribution in Figure 2. Regarding this distribution of optimal threshold values from all typically-developing participants, the mode was 0.5, the median was 7.1, the mean was 14.67, and the range was 0 – 64.25. Based on the right skew of this distribution and the presence of a few high-value outliers, it was determined that the median (7.1) would be the most appropriate measure of central tendency. Therefore, this median optimal threshold value was applied to data from all participants, to calculate the individual sensitivity, specificity, and overall agreement levels at this threshold of 7.1.

Figure 2. Frequency Distribution of Accelerometry-Based Optimal Thresholds – TD Group

Abbreviations: TD, typically-developing.
4.3.2 Sensitivity, Specificity and Youden’s Index at Different Thresholds

Sensitivity and specificity values (and Youden’s Index) were calculated for each participant’s data (pooled for dominant and non-dominant arms) at the median optimal threshold value of 7.1. These values are presented in Figure 3, with a comparison to the values at the (default) zero threshold, and a threshold of 2.0. As stated in the Methods section, sensitivity, specificity values and Youden’s Index were calculated at a threshold of 2.0 to allow for a comparison of results to those obtained by studies that validated arm accelerometry for use in post-stroke adults and investigated arm accelerometry in children.50,52,53

<table>
<thead>
<tr>
<th></th>
<th>At Zero</th>
<th>At 7.1</th>
<th>At 2.0</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sensitivity</td>
<td>86.44%</td>
<td>88.43%</td>
<td>87.85%</td>
</tr>
<tr>
<td>Specificity</td>
<td>74.16%</td>
<td>77.85%</td>
<td>75.51%</td>
</tr>
<tr>
<td>Youden's Index</td>
<td>62.98%</td>
<td>64.33%</td>
<td>63.95%</td>
</tr>
</tbody>
</table>

Figure 3. Sensitivity, Specificity and Youden's Index Values – TD Group

Abbreviations: Error bars indicate SD values for each distribution of data.

As shown Figure 3, mean sensitivity was the highest at a zero threshold and mean specificity was highest at a threshold of 7.1. Youden’s Index, which takes into account both sensitivity and specificity with equal weighting, was similar across the three threshold settings (ranging from 62.98 – 64.33%). Specifically, based on the results of a 1-way ANOVA, no significant differences were found in Youden’s Index values between the three thresholds (F = 0.133, df 2, p-value = 0.876).
4.3.3 Observer vs. Accelerometer Agreement Levels at Different Thresholds

In continuing to address Objective 1, observer and accelerometer agreement levels were calculated at the median optimal threshold for the entire group (7.1), and compared to these values at a zero threshold and a threshold of 2.0 (see Figure 4).

Figure 4. Comparing Agreement Levels at Thresholds of Zero, 7.1 and 2.0 – TD Group

Abbreviations: Error bars indicate SD values for each distribution of data.

As displayed in Figure 4, mean accuracy levels were similar across the threshold settings of 0, 2.0, and 7.1 for dominant arm (ranging from 86.56 – 88.44%) and non-dominant arm (ranging from 83.31 – 84.43%) movements. Based on the results of a 2-way ANOVA (repeated measures), it was found that agreement levels were significantly higher for dominant arm compared to non-dominant arm movements, at all three thresholds (F=27.980, df 1, p-value < .0001).

4.3.4 PABAK Statistic – Accelerometer-Observer Agreement at Different Thresholds

PABAK values, also indicating accelerometer-observer agreement, for thresholds of zero, 7.1 and 2.0 are shown in Figure 5.
At the different thresholds tested, PABAK coefficients were similar, differing by 0.03 at most (between thresholds of zero and 7.1). Based on the rating scale (Table 2) used for interpreting this statistic, PABAK coefficients at all thresholds tested (zero, 7.1, and 2.0) were within the Good range.

4.4 Arm Use Ratios in Typically-developing Group

Arm use ratios and further analyses of TD data were calculated based on a 2.0 threshold setting. This threshold setting was chosen for two reasons. First, Youden’s Index and the level of agreement were similar across the three threshold settings investigated. Second, previous published studies of arm accelerometry (for adults and children)\(^{50,52,53}\) used a threshold of 2.0, thus selecting this threshold setting allowed for comparison of the results to previous work.

4.4.1 Arm Use Ratios – Comparison between age groups

Accelerometry-based arm movement ratios were calculated for each TD participant. These arm ratios were compared between different age groups of typically-developing participants, as shown in Figure 6.
The youngest (3-year old) age group had the highest mean arm use ratio (93.41 +/- 4.46%) and the oldest (8 – 10-year old) age group had the lowest mean arm use ratio (83.21 +/- 8.66%). However, differences between the mean arm use ratios for all age groups were found to be non-significant, based on an AVOVA (F = 2.127, df 3, p-value = 0.126).

4.5 Results for Group with Hemiparesis

4.5.1 Descriptive and Clinical Measures for the Group with Hemiparesis

For participants with hemiparesis (n = 13), their descriptive and clinical measures related to arm movement and function are presented in Table 5. Based on their QUEST (Dissociated Movement and Grasp Scales) and PMAL-R Scores, 12 out of the 13 participants were estimated to have some level of affected arm disregard (DDI > 0). For 11 out of 13 participants, at least some mirror movements were observed during the fine-motor screening (MM Score > 0). The MACS Level ratings for all participants ranged from I – II, with the majority of participants (8/13 or 61.54% of the sample) being classified at Level I.
Table 5. Descriptive and Clinical Measures for Participants with Hemiparesis

<table>
<thead>
<tr>
<th>Participant #</th>
<th>QUEST Sub-Scores</th>
<th>PMAL-R Scores:</th>
<th>Disregard Index (DDI)</th>
<th>Mirror Movement Scores:</th>
<th>MACS Movement Level</th>
<th>MTS Wrist, Elbow</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>D: 92.0, 84.4</td>
<td>HO: 1.09</td>
<td>70.2</td>
<td>Raw: 7</td>
<td>I</td>
<td>0.2</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>D: 100, 100</td>
<td>HO: 4.73</td>
<td>-7.1</td>
<td>Raw: 0</td>
<td>I</td>
<td>0.0</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>D: 60.9, 40.6</td>
<td>HO: 1.09</td>
<td>39.14</td>
<td>Raw: 4</td>
<td>II</td>
<td>2.2</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>D: 100, 100</td>
<td>HO: 1.75</td>
<td>65.0</td>
<td>Raw: 3</td>
<td>I</td>
<td>0.0</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5+</td>
<td>D: 78.9, 75.0</td>
<td>HO: 2.0</td>
<td>38.91</td>
<td>Raw: 0*</td>
<td>I</td>
<td>NT</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6+</td>
<td>D: 56.2, 53.1</td>
<td>HO: 0.68</td>
<td>42.64</td>
<td>Raw: 12</td>
<td>II</td>
<td>1.2</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>D: 81.2, 46.9</td>
<td>HO: 2.59</td>
<td>29.45</td>
<td>Raw: 3.5</td>
<td>I</td>
<td>2.0</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>D: 67.9, 50.0</td>
<td>HO: 2.23</td>
<td>23.36</td>
<td>Raw: 3</td>
<td>II</td>
<td>2.2</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>D: 86.7, 42.2</td>
<td>HO: 2.36</td>
<td>39.52</td>
<td>Raw: 1.5</td>
<td>II</td>
<td>NT</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Avg (0-4): 1.17

II: 2.2

Average: 0.67

Avg: 0.5
<table>
<thead>
<tr>
<th></th>
<th>D: 76.6,46.9</th>
<th>HO: 2.32</th>
<th>30.16</th>
<th>Raw: 3</th>
<th>I</th>
<th>2,2</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>G: 72.2,48.1</td>
<td>HW: 2.27</td>
<td></td>
<td>Avg:0.5</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

| 11| D: 61.7,23.4 | HO: 1.98 | 22.12 | Raw: 8 | II | 3,3 |
|   | G: 51.9,33.3 | HW: 2.09 |       | Avg: 1.33 |   |     |
|   |              |         |       |        |   |     |
| 12| D: 93.0,73.4 | HO: 2.45 | 44.0  | Raw: 7.5 | I | 1.1 |
|   | G: 59.3,55.6 | HW: 2.27 |       | Avg: 1.25 |   |     |
|   |              |         |       |        |   |     |
| 13| D: 90.6,87.5 | HO: 2.59 | 2.62  | Raw: 9  | I | 1.1 |
|   | G: 86.1,96.3 | HW: 2.36 |       | Avg: 1.5 |   |     |

Abbreviations: D, Dissociated Movement Score; G, Grasp Score; HO, How Often Scale; HW, How Well Scale; Avg, Average; MTS, Modified Tardieu Scale; Bilat, Bilateral; Uni, Unilateral; NT, Not Tested (due to the difficulty with these participants in assuming a relaxed arm posture for passive stretching). * Motor overflow arm movements were observed in this participant during the fine-motor screening, despite the lack of Mirror Movements. + These two participants had a concurrent diagnosis of Autism Spectrum Disorder (ASD) as well as hemiplegic CP.

4.6 Accelerometer Accuracy – Group with Hemiparetic Cerebral Palsy (CP)

As was performed for the TD group, the sensitivity and specificity of accelerometry-based measurements were also calculated for the CP group. Initial sensitivity and specificity values were determined at a default VM threshold setting of zero. At this zero threshold, mean sensitivity values (84.8 +/- 7.24%) were slightly higher than mean specificity values (84.0 +/- 8.96%), for data from both dominant and hemiparetic arms, for all participants in the CP group.

4.6.1 ROC Analysis: Sensitivity and Specificity at Different Threshold Settings

As performed for the TD group, ROC Analysis and the calculation of Youden’s Index were performed for the CP group data.
4.6.2 Threshold for Optimal Sensitivity and Specificity

ROC analysis was performed on each data set from each participant with hemiparesis, and the optimal threshold setting was determined based on Youden’s Index. Figure 7 presents a frequency distribution of the optimal threshold values for the CP group. As shown in Figure 9, optimal threshold (VM) values between 0 and 5 were found with the greatest frequency in this group, as was found for the TD group.

![Figure 7. Frequency Distribution of Accelerometry-Based Optimal Threshold – CP Group](image)

Abbreviations: VM, vector magnitude.

Based on this frequency distribution, the mode was 0.5, the mean optimal threshold value was 9.1, the median value was 4.88, and the range was 0 - 66.03. Based on the right skew of this distribution and the presence of a few (higher value) outliers, it was determined that the median value (4.88) would be the best estimator of central tendency.

4.6.3 Sensitivity, Specificity and Youden’s Index at Different Thresholds

Sensitivity, specificity and Youden’s Index were calculated for each CP participant’s data at the median optimal threshold of 4.88. These results are displayed in Figure 8, with a comparison to the same values at thresholds of zero and 2.0.
As shown in Figure 8, mean sensitivity was highest at the zero threshold and mean specificity was highest at a threshold of 4.88. As an index that weighs both sensitivity and specificity equally, Youden’s Index appeared to be highest at the median optimal threshold of 4.88. At a threshold of 2.0, the mean Youden’s Index value appeared to be slightly higher than the mean at a zero threshold. However, differences in Youden’s Index between the three thresholds were not significantly different, based on the results of a 1-way ANOVA ($F = 0.150$, df 2, p-value = 0.861).

### 4.6.4 Observer vs. Accelerometer Agreement at Different Thresholds – CP Group

In continuing to address Objective 1, observer/accelerometer agreement levels for data from participants in the CP group were calculated at the median optimal threshold for the entire group (4.88), and compared to these values at thresholds of zero and 2.0 (see Figure 9).
As shown in Figure 9, mean accuracy levels, averaged for both arms, were similar across the three threshold settings tested, with higher accuracy for dominant compared to hemiparetic arm movements. Based on the results of a 2-way ANOVA, differences between accuracy for dominant and non-dominant arm movements were statistically significant across all three thresholds ($F = 7203.408$, df 2, $p < 0.0001$).

PABAK values for the three threshold settings of zero, 4.88 and 2.0 are shown in Figure 10.
At the different thresholds tested, PABAK coefficients differed at most by 0.029, between thresholds zero and 4.88, for hemiparetic arm movements. At thresholds 2.0 and zero, PABAK coefficients (for both arms) were similar, differing by only .012. Based on the rating scale (Table 2) used for interpreting these values, PABAK coefficients at all thresholds tested (zero, 4.88 and 2.0) were within the Good range.

Based on these results, use of a 2.0 VM threshold setting for subsequent analyses of accelerometry data from the CP group (as well as TD group) was used.

### 4.7 Correlational Analysis: Clinical Scores vs. Arm Use Ratios – CP Group

In addressing Objective 2, participants’ (with hemiparesis) accelerometry-based arm use ratios were compared, via correlational analysis, to their scores on the QUEST (i.e. Dissociated Movement and Grasp scale scores).

#### 4.7.1 QUEST Scores vs. Accelerometry-based Arm Use Ratios

A bivariate correlational analysis of accelerometry-based arm use ratios and QUEST (Dissociated Movement scale) scores for all 13 participants with hemiparesis indicated no significant
correlation (bilateral QUEST scores: $r_s = 0.006$, $p = 0.986$, unilateral QUEST scores: $r_s = 0.116$, $p = 0.707$).

The bivariate correlational analysis was repeated with the two participants with a concomitant diagnosis of ASD removed. For participants with only a CP diagnosis (n = 11), this correlation was also found to be non-significant (bilateral QUEST scores: $r_s = 0.223$, $p = 0.509$, unilateral QUEST scores: $r_s = 0.187$, $p = 0.581$).

Based on observations during videotaped play that child participants presented with very small arm movements that were often non-functional and appearing to be involuntary (i.e. during times when they were instructed to sit still), these correlational analyses were repeated at a higher accelerometry-based threshold setting, excluding movements at or below the average VM that corresponded to observations of very small movements (see Methods, Chapter 2 for a description of how this threshold was calculated). At this higher threshold, a moderate correlation was found between arm use ratios and QUEST (Dissociated Movement scale) scores, for participants with only CP (n = 11). When bilateral QUEST scores were used for this analysis, the correlation coefficient was: $r_s = 0.592$ ($p = 0.055$), on the borderline of significance at an alpha level of 0.05. When unilateral (hemiparetic) QUEST scores were used for this analysis, the correlation coefficient was: $r_s = 0.662$, $p = 0.026$, significant at an alpha level of 0.05. Figure 11 displays a scatter plot with the data points for each participant included in this correlational analysis, for which a statistically significant relationship (at an alpha level of 0.05) was found.
Figure 11. Unilateral QUEST (Dissociated Movement) Scores and Arm Use Ratios (>vs)

Abbreviations: QUEST, Quality of Upper Extremity Skills Test; Diss, Dissociated Movement Scale; vs, very small, $r_s$, Spearman’s rank correlation coefficient; p, p-value. * indicates a statistically significant result at an alpha level of 0.05.

No significant correlations were found between QUEST (Grasp Scale) scores and participants’ arm ratios ($p>0.05$) when calculated under all of the aforementioned threshold and participant inclusion conditions.

Specifically, at a threshold excluding very small movements (>vs), including participants with only CP ($n=11$), the level of correlation between bilateral QUEST (Grasp scores) and arm ratios was: $r_s = 0.323$ ($p = 0.332$) and between unilateral QUEST (Grasp scores) and arm ratios was: $r_s = 0.464$ ($p = 0.151$).
4.7.2 PMAL-R Scores vs. Accelerometry-based Arm Ratios

A bivariate correlational analysis of accelerometry-based arm ratios (calculated at a threshold of 2.0) and PMAL-R (How Often) scores for all participants with hemiparesis indicated no statistically significant relationship, $r_s = 0.179$, $p = 0.558$.

As well, when this correlational analysis was performed on arm ratios calculated at a higher threshold (including > very small movements) for participants with only a CP diagnosis ($n = 11$), no statistically significant relationship was found ($r_s = -0.037$, $p = 0.915$). Figure 12 displays a scatter plot with data points for each participant included in this analysis.

![Figure 12. PMAL-R (HO) Scores and Arm Use Ratios (>vs) for Participants with CP](image)

**Figure 12. PMAL-R (HO) Scores and Arm Use Ratios (>vs) for Participants with CP**

*Abbreviations: PMAL-R, Paediatric Motor Activity Log-Revised; HO, How Often Scale; vs, very small arm movements; $r_s$, Spearman’s rank coefficient; p, p-value.*

4.7.3 Additional Exploratory Analyses

Based on the descriptive measures collected for participants with hemiparesis, including observed Mirror Movement (MM) scores, and Disregard Index (DDI) scores, exploratory analyses were performed to determine if any correlations could be found between these variables and the accelerometry-based arm use ratios. However, no statistically significant correlations were found
between accelerometry-based arm use ratios and MM scores (p > 0.339) or DDI scores (p > 0.401).

4.8 Arm Movement Ratios: TD Group vs. Group with CP

In addressing Objective 3, accelerometry-based arm movement ratios were compared between a set of nine age- and sex-matched pairs from the TD and CP groups, respectively.

4.8.1 Arm Movement Ratios At an Accelerometry Threshold of 2.0

In comparing arm movement ratios calculated at an accelerometry threshold of 2.0 for all nine matched pairs, the mean paired difference in arm ratios was 9.92 (SD = 8.95), with TD participants’ arm use ratios being greater, overall, than those of participants in the CP group. This difference was found to be statistically significant based on a paired t-test (t = -3.324, df 8, p = 0.01). The arm movement ratios for each participant in this analysis are displayed in Figure 13.

![Figure 13. Arm Ratios of Matched Pairs from TD and CP Groups](image)

Abbreviations: CP, cerebral palsy; TD, typically developing; VM, vector magnitude.
As shown in the above Figure, arm ratios (calculated at a threshold of VM = 2.0) were lower for eight out of nine participants with hemiparetic CP, compared to a typically developing age- and sex-matched participant.

4.8.2 Arm Movement Ratios At a Threshold Excluding Very Small Movements

In comparing arm movement ratios at a higher threshold (excluding very small movements), these ratios were also lower for eight out of nine participants with hemiparetic CP compared to the respective match’s ratio in the TD group, as shown in Figure 14. The mean paired difference in these arm movement ratios was 29.234 (SD = 19.2), and this difference was also found to be statistically significant, based on a paired t-test (t = -4.567, p = 0.02).

Figure 14. Arm Ratios (for > vs movements) of Matched Pairs from CP and TD Groups

Abbreviations: CP, cerebral palsy; TD, typically developing; vs, very small.
Finally, Table 6 displays a summary of the accelerometry-based arm ratios found for the matched participants from the CP and TD groups. Arm ratio values for these nine matched pairs were similar to the mean ratios calculated for the entire group (i.e. at a threshold of 2.0, the mean arm ratio for the entire TD group (n = 26) was 0.885 +/- 6.9 and for all participants in the CP group (n = 13), it was 0.778 +/- 6.2).

| Table 6. Mean Accelerometry-based Arm Ratios for Matched Pairs (n = 9/group) |
|-------------------------------------------------|-----------------|-----------------|
| CP group                                        | TD group        |
| At threshold 2.0                                | 0.785 (+/- 6.7) | 0.885 (+/- 7.9) |
| At threshold > vs movements                     | 0.496 (+/- 15.2) | 0.773 (+/- 14.2) |

Abbreviations: CP, cerebral palsy; TD, typically-developing; vs, very small.
Chapter 5

5 Discussion

Overview of Discussion

This study investigated the validity of arm-accelerometry, as a measure of arm movement, in children with and without hemiparesis. Three types of validity were evaluated here: criterion (concurrent), convergent and discriminative. The results indicated that arm-accelerometry has criterion (concurrent) validity in detecting the presence/absence of arm movements in both TD and CP groups, based on sensitivity, specificity (Youden’s Index, $J > 0.63$) and agreement ( > 84%, PABAK > 0.69), with reference to a criterion of video-based observations. The inter-rater reliability of these video-based observations was also determined to be in the ‘Good’ range (PABAK > 0.79) for determining the presence of arm movements. Regarding the convergent validity analysis, accelerometry-based arm ratios were found to moderately correlate with unilateral QUEST scores (Dissociated Movement) of $n=11$ participants with only a CP diagnosis. However, this accelerometry-based index (i.e. arm ratios) was not found to correlate with the QUEST Grasp scores or PMAL-R scores of these participants. With respect to discriminative validity, accelerometry-based arm ratios were found to differ significantly between children with and without hemiplegic CP, based on a sample of 9 pairs, matched in age and sex.

This chapter will include a consideration of these results, in relation to findings from other published work. Major limitations of this research project will also be identified. Finally, the implications of these research findings and future directions for research will be highlighted.

5.1 Inter-rater reliability testing

Regarding inter-rater reliability testing of video-based observations, a Good level of agreement (PABAK > 0.79) was found between observers (with varying academic/professional backgrounds) in identifying the presence of arm movements. Thus, it can be expected that the video analysis, as performed here, would be transferable to different researchers in this setting, regardless of their research/clinical background. The observational scoring results obtained here do not appear to be dependent on or biased by the ratings of one specific researcher.
However, lower agreement was found between observers regarding the size of observed arm movements (.52 < PABAK < .72). This did not affect the validity of data used for addressing the main objectives of this study. This is because the majority of analyses performed here were based on observations of the presence/absence of movement. Observations of ‘very small’ movements were only used in conjunction with accelerometry-based VMs, to inform the determination of a cut-off threshold point that would exclude these movements, in addressing Objective 2. However, the discrepancy between observers in rating the size of observed arm movements does indicate the need for a more objective, standardized method of systematically assigning these size scores.

As an example of a more systematic approach to scoring arm movement size, Uswatte et al. (2000) used a scale that assigned discrete measurements (i.e. < 2.5 cm forearm displacement or < 30° supination/pronation = no movement; 2.5 – 7.6 cm or 30° < supination/pronation < 45° = small movement; > 7.6 cm and > 45° = large movement) of forearm movements on the video screen to different size categories. Use of this approach for future research, involving an analysis of arm movement size, could increase the inter-rater reliability when measuring this construct.

Such a systematic, objectively quantifiable approach will be important to establish for future analyses. It is noteworthy that a significant relationship between accelerometry-based arm ratios and unilateral QUEST scores was only found when the average level of movements observed to be ‘very small’ were excluded. This finding reinforces the need for accuracy and reliability when scoring movements as ‘very small’, so that consistently high thresholds can be set that will allow for the exclusion of these very small movements from all child participants in a systematic manner.

5.2 Determining an Optimal Threshold Setting for Accelerometer Vector Magnitudes

Previous research that validated arm-accelerometry as a measure of arm use in post-stroke adults found 2.0 to be an optimal threshold setting, based on their comparison of accelerometer-observer agreement levels at different thresholds, for n = 8 participants.53 In using this setting, all movements with an accelerometry-based vector magnitude (VM) less than 2.0 were considered to be ‘no movement’. Therefore, VM recordings greater than or equal to 2.0 were considered to indicate ‘movement’. This threshold of 2.0 was also used by Sokal et al. (2016) in their study.
involving child participants, however no justification for use of this threshold with children was provided. The current study thus undertook a comparative investigation of different threshold settings, and the effects on accelerometry’s sensitivity, specificity and accuracy levels (i.e. % agreement and PABAK coefficients) in measuring arm movements in children. This analysis revealed that i) optimal thresholds (based on ROC analysis and Youden’s Index) were variable between participants (ranging from 0 – 64.5), with the majority of optimal thresholds in the range of 0 – 5, in both TD and CP groups; and ii) applying the median optimal threshold for each group (7.1 for TD group, 4.88 for CP group) yielded similar sensitivity, specificity and accuracy results to use of thresholds 0 and 2.0. Although the mean Youden’s Index appeared to be slightly higher at a threshold of 2.0 compared to a zero threshold for both TD and CP groups, these differences were not significant. These results collectively led to the selection of use of 2.0 as the threshold for this project’s subsequent analyses, as it allowed for a comparison of the current findings to those of Uswatte et al. (i.e. the primary validation study of arm accelerometry in adults) and Sokal et al. (i.e. a study exploring arm accelerometry’s use in children with CP). Regarding the identification of a threshold for future research and clinical applications with children, the 2.0 threshold appears to be a more feasible choice based on this current study’s results. This is because there is similar accuracy, sensitivity and specificity at this threshold, compared to the use of the median optimal threshold found for each group tested here. As well, a threshold of 2.0 can be applied efficiently to a whole group, without requiring future individualized ROC analysis of each participant’s data.

The current finding of high sensitivity, specificity and agreement at a threshold of 2.0 indicates that the optimal threshold for groups of children (TD and CP) tested here is not different than the optimal threshold found for adults by Uswatte et al. (2000).

5.3 Accelerometer Accuracy

Accelerometry’s accuracy in measuring children’s arm movements was important to establish, as no published results were found that validated this measurement tool’s use in typically-developing children. Although one study examined arm-accelerometry’s convergent validity (with the PMAL-R and PAFT) in children with CP, they did not perform any criterion validity testing of this tool in children with CP or in a typically-developing group.50
The observer-accelerometer agreement found here for detection of arm movements was good overall (> 84%, PABAK > 0.69) and was similar for TD and CP groups, supporting this tool’s potential for use with clinical paediatric populations.

Although levels of observer-accelerometer agreement were adequate in validating this measure, they were lower than the agreement levels found in published literature for post-stroke adults, which ranged from 0.93\textsuperscript{53} to 0.96\textsuperscript{54}.

One possible reason for this discrepancy is that there is a higher preponderance of non-functional (and often very small) movements in children, compared to adults, and that this makes observer scoring for this population more difficult. Sokal et al.’s (2016) findings are consistent with this possibility, as they deduced that the child participants with CP (n = 21) in their study were presenting with a high level of movement that was non-functional when wearing accelerometers in natural environments. Specifically, Sokal et al. reported that, based on visual inspection of their data and the correlations found between accelerometry measures, their results were consistent with the hypotheses that children (in their study) both engaged in more bimanual activities and moved their more-affected arm even when using the less-affected arm alone, to a greater extent than adults (in previous studies). The latter hypothesis would be consistent with a higher preponderance of non-functional arm movements in children compared to adults.

Another reason for this discrepancy in agreement levels could be the different methodology used by Uswatte et al. in the adult studies. As previously described, Uswatte et al.’s approach involved having observers disregard (i.e. assigning a score of zero) any arm movements that were equal to or less than 2.5 cm in displacement length or < 30° supination/pronation). This approach involves one method of eliminating the presence of very small movements at the video-observation stage of data collection. Uswatte et al. also found that the highest level of disagreement between accelerometer and observer was found in movements that were observed to be smaller (as opposed to larger), based on their coding system. Specifically, they found that accelerometer-observer agreement was 98% when ‘small’ movements were excluded from analysis. However, agreement was only 88.7% when ‘small’ movements were included. Based on Uswatte et al.’s results, it appears that the inclusion of relatively smaller movements in agreement analyses contributes to a decrease in agreement levels. Thus, the lower agreement levels found here (compared to Uswatte
et al.’s adult study) could in-part be explained by the inclusion of very small and small movements in the observers’ scoring of arm movements used in this study.

As well, Uswatte’s use of discrete (cm) measurements for observers to base their movement scores on, was a more objective and quantifiable approach to scoring than was used here. Use of their more systematic approach could have increased the level of agreement between observers and accelerometer scores, relative to this study’s findings.

5.4 Accelerometry-based Arm Movement Ratios

In this study, arm ratios were initially measured at an accelerometry-based threshold of 2.0, in typically developing (TD) children (n = 26) and children with hemiparetic CP (n = 13). Arm ratios were also compared between nine age- and sex-matched pairs of children with and without hemiplegic CP, at both a threshold of 2.0, and at a threshold that excluded very small movements (see Methods section for an explanation of how this was calculated).

Regarding arm ratios in individuals with hemiparesis, the mean arm ratio found here for children with hemiplegic CP was 0.78 (at a threshold of 2.0) and 0.50 (at a higher threshold excluding very small movements). Compared to the arm ratios of 0.86 found by Sokal et al. (2016) for children with CP (also using a threshold of 2.0), the mean ratios found here were lower by 8%. It is possible that the differences in time periods of measurement led to different ratio results. Sokal et al. measured arm use in children over an extended period (> 9 hours/day, for 3 days), however this study involved measurement over a much more condensed period of 20 minutes. It is possible that arm ratios would have been greater if the measurement period was extended to that length of time. This expectation is logical based on the assumption that over a longer period of time in natural environments, more rest time (i.e. sitting, lying) would be included. During rest times, there would be a greater likelihood that the arms would be approximately doing the same thing (i.e. both not moving or moving minimally), thus increasing the non-dominant:dominant arm movement ratio.

Considering both the hemiparetic arm ratios found in this study and Sokal’s (0.78-0.86) to those published for hemiparetic adults, children with hemiparesis appear to have higher arm ratios than those found for post-stroke adults. Specifically, published adult studies of post-stroke adults with hemiparesis report arm ratios ranging from 0.45 – 0.58. These results are consistent with
the possibility of a higher preponderance of non-functional (affected) movements and/or a higher level of bimanual activity, in children compared to adults, with hemiparesis.

However, for the typically-developing children, the mean arm ratio found here (for \( n = 26 \)) was 0.885 (\( \pm 6.9 \)), which is within the range found for healthy adults by other studies. Specifically, published reports of healthy adult arm ratios range from 0.79 to 1.1. \(^{72}\) Bailey & Lang (2014) published referent arm ratio values for healthy adults (based on accelerometry data over a 24-hour period) of 0.95 (\( \pm 0.06 \)). Compared to the mean ratio found here for healthy children (0.885 \( \pm 6.9 \)), it cannot be assumed that healthy children have a higher arm ratio than adults. However, the difference in measurement time between this study (20-minute period) and Bailey & Lang’s (24-hour period) as well as that of other adult studies (ranging from 1 – 5 day periods of data collection) is substantial, and it is recommended that arm ratios for healthy children be measured over longer periods of time, to allow for equal comparisons to be made between child and adult arm ratios. This study is unique in its findings of arm ratios for typically-developing children. There is no published research yet on typically-developing children’s arm use ratios, and it will be important for future research to provide such normative data as a basis for comparison to arm ratios in clinical paediatric populations.

For this study, summary variables of arm movement (i.e. the proportion of epochs that had VMs above a threshold value) were used to calculate arm ratios, as done in most studies for adults. Additionally, a few studies also looked at intensity variables (i.e. the mean VM value for all epochs, divided by the number of epochs). This could be explored in future analyses with the data, to see if intensity-based arm ratios differ between TD and CP individuals, and if they correlate more with convergent measures of arm use.

For the typically-developing group of children tested here, their arm movement ratios did not differ significantly between age groups. However, based on the higher level of bilateral symmetrical arm use in younger children compared to older children (who have more differentiated handedness and isolated dominant arm use)\(^{76}\), it would be expected that children’s arm movement ratios would decrease with developmental age.

The lack of significant differences found here could be explained by the small group sizes (i.e. \( n = 5 - 9 \) per group). It is recommended that this analysis be performed in future, with greater sample
sizes per group. It is also possible that differences in arm ratios between different age groups (of TD children) could be revealed with the use of a higher threshold setting, that excludes very small or smaller movements. The use of this higher threshold setting may involve a compromise in the level of sensitivity, specificity and accuracy of accelerometry-based measures. However, by selecting larger movements (that are potentially more functional and purposeful), these arm ratios could perhaps better reflect more functional arm use and differences between younger and older children. This is another investigation that is recommended for future analysis. It will also be useful to have normative values for arm use ratios in typically-developing children, as a reference for paediatric clinical work, as no referent values have been published for healthy children, to date. Establishing these referent values for healthy children of different ages could provide norms, for comparison to arm ratios in children with arm impairments (such as hemiparesis), that could lead to better estimation of the degree of arm impairment.

5.5 Rationale for the Decision to Eliminate Very Small Movements

The reason for the decision to calculate arm ratios at a higher threshold excluding very small movements was based on video observations (during data collection) that the majority of very small observed movements were not contributing to functional/purposeful movements and appeared to be often involuntary (e.g. arm movements during still periods).

In the adult arm accelerometry validation study, Uswatte et al. asked observers to disregard very small movements observed (i.e. < 2.5 cm) in adults and perhaps they had the same rationale as stated above. Also, based on Sokal et al.’s postulation that children made more arm movements that were non-functional than adults, it was considered that perhaps arm ratios at a higher threshold would better capture information related to actual/functional arm use than arm ratios at a lower threshold. The potential value of using a higher threshold was supported here based on the finding of a significant correlation between arm use ratios and unilateral QUEST (Dissociated Movement) at this threshold (convergent validity analysis) and a significant difference between CP and TD participants’ arm ratios found at this threshold (discriminative validity analysis). Further tests of accelerometry’s accuracy, sensitivity and specificity at this higher threshold (excluding very small movements) are recommended for future analysis.
5.6 Convergent Validity Analysis

Although a moderate significant correlation was found between accelerometry-based arm ratios and hemiparetic QUEST (Dissociated Movement) scores, no significant correlation was found with the QUEST (Grasp Scale) scores. This can be explained by the fact that Dissociated Movement scores target the ability to produce isolated movements at upper extremity joints, regardless of the child’s dexterity with functional tasks (which the Grasp score targets). Thus, the Dissociated Movement scale measures a construct that is more similar to what is measured by the accelerometers, as they capture quantitative information about movement only, regardless of whether it is functional or dexterous.

Regarding the QUEST (and other clinical screens) scoring procedures used here, all scoring was completed by one OT (J. Dawe). Therefore, inter-rater differences in QUEST (and other clinical screens) did not need to be accounted for in this study. Although this researcher received training in administering and scoring the QUEST, the study design could be modified in future by having the QUEST administration videotaped, with scoring checked by other independent clinician observers. However, videotaping of the QUEST is not standard practice for the administration of this assessment.

With respect to the QUEST’s use in this study, its validity was only established for a younger age range (18 months – 8 years) than was tested here (CP group: age 4 – 14 years). Although the inter-rater reliability of the QUEST was established for children up to 16 years of age\(^73\) (with Acquired Brain Injury) and for children as old as 12 years\(^74\) (with CP), its validity has not been established for ages older than 8 years. Therefore, the validity of the QUEST scores for the older participants (ages 9 – 14) in this sample (5/11 participants with only a CP diagnosis) cannot be confirmed. This limits the validity of correlational analyses between arm ratios and QUEST scores, as a means to establish arm-accelerometry’s convergent validity.

Accelerometry-based arm ratios, as measured here, did not correlate with children’s PMAL-R scores. There are several possible reasons for this lack of correlation. First, the PMAL-R is a subjective measure, based on parent-report and not based on objective quantification of arm movement, as accelerometry data is. Limitations in PMAL-R’s validity include that it is subject to
parental recall bias, and its moderate convergent validity has only been established for measuring change scores from pre- to post-treatment in children ages 2 – 6 years\textsuperscript{75}, and not for measuring arm use in older children or based on a one-time administration of the questionnaire. Furthermore, the PMAL-R was found to have high between-subject variability when administered only once\textsuperscript{39} (as it was used in this study). Thus, there is not enough evidence to ascertain the validity of PMAL-R scores obtained here for all participants with hemiparesis.

Sokal et al. (2016) also found a low correlation between PMAL-R and accelerometry-based scores in children with CP, however they also based their analysis on a one-time administration of the PMAL-R, for which this test has not been validated. As further evidence of PMAL-R’s questionability as a valid paediatric outcome measure (based on 1-time administration), Sokal et al. also did not find a strong correlation between PMAL-R and the other measure of paediatric arm function in their study, the PAFT. In fact, some of the accelerometry-based data that they collected correlated more strongly ($r_s = 0.52 - 0.55$, $p = 0.003 - 0.007$) with the children’s PAFT scores than the PMAL-R scores did with the PAFT ($r_s = 0.43$, $p = 0.03$). Therefore, Sokal’s conclusion that arm accelerometry may not be a useful index of rehabilitative real-world outcome for children, based on a low correlation with PMAL-R scores, does not appear to be justified.

Regarding the other exploratory correlational analyses that were investigated here, no significant correlations were found between arm ratios and children’s DDI scores or Mirror Movement scores. Although no formal hypotheses were made regarding these measures, it could be expected that DDI scores (estimating the level of affected arm disregard in daily life) would be negatively correlated with arm ratios. However, the fact that the DDI score is a composite based on both the QUEST and PMAL-R scores, indicates that the threats to validity of PMAL-R scores would equally threaten the validity of DDI scores. Another possible expectation could be that mirror movement scores (indicating a greater degree of bilateral symmetrical arm movements, despite the unintended nature of them) would be positively correlated with arm ratios. However, such a correlation was not found here. This lack of a significant finding could be due to the small sample size of children with a CP only diagnosis ($n = 11$) that were examined for this analysis. Based on sample size calculations (with beta = 0.20 and alpha = 0.05), a sample size of $> 19$ would be required to have 80% power to detect a correlation between two continuous variables with a correlation of $r = 0.6$ or higher. Thus, for non-continuous ordinal variables (i.e. DDI and MM
scores), the projected sample size required to demonstrate a significant result with 80% power would be even higher than \( n = 19 \). Thus, it is recommended that the potential correlation between Mirror Movement scores and accelerometry-based arm ratios be re-investigated with a larger sample of children with hemiplegic CP.

5.7 Discriminative Validity Testing

At both a threshold setting of 2.0 and at a higher threshold (excluding very small movements), arm ratios were found to significantly differ between children with and without hemiplegic CP (based on 9 pairs, matched in age and sex).

With the exclusion of very small movements, a greater magnitude of difference was found between CP and TD matches in arm use ratios than at a lower threshold of 2.0 (although the significance level was lower, and almost equal). This finding of a great magnitude of difference in arm ratios (that is statistically significant) at a high threshold reinforces the potential value in using this higher threshold. Thus, future research should include both the establishment of a systematic method for eliminating very small movements and a formal criterion validity test of accelerometry’s accuracy at this higher threshold. To avoid the design of a validity test that would be self-referential, it is recommended that another method of objectively measuring the size of observed arm movements be used for this purpose (as a criterion reference), such as the metrics obtainable from 3D motion analysis or Kinect-based videotaping.

Based on the small sample size (\( n=11 \)) of children with CP, and the narrow range of their arm use abilities, between-subject comparisons of arm ratios of children with CP could not be analyzed. However, with use of larger sample sizes, it will be worthwhile to investigate if arm ratios differ significantly between groups of children with CP who are classified at different levels of upper extremity function (i.e. based on their QUEST or other clinical scores).

5.8 Limitations of Study

A major limitation of this study is the small sample size of participants with a CP diagnosis (\( n=11 \)) and the small sample size of matched-pairs between the CP and TD groups (\( n=9 \) pairs). As stated previously, the small sample size limits the power of this study to detect significant correlations between clinical scores and accelerometry-based arm ratios.
Another limitation is the unnatural controlled lab setting that was used for observational and accelerometry-based data collection. This controlled setting may have increased the internal validity, by controlling for differences between participants in the pattern of wear and the activities performed. However, the unnatural setting limited the external validity of this study’s findings. We do not have enough information yet to generalize these findings to natural environments of participants. Thus, further validity studies of accelerometry in children’s natural settings are required.

In addition, the use of video data (that was recorded based on a 2-dimensional frontal view of each participant) for observer-based scoring of arm movements could have limited the ability to accurately monitor movement outside of the visible planes. For example, based on a frontal view of a participant, it is possible that he/she could have produced movement(s) in the sagittal plane (e.g. shoulder extension) that were not directly visible to the observer viewing the videotape recording. This limitation could have led to a tendency for observation-based scores of movement to under-represent the amount of movement actually occurring, and could have contributed to any observer vs. accelerometer disagreement.

As well, there is a need for more precision in rating arm movement size, so that very small arm movements can be systematically identified, and an optimal high accelerometry-based threshold setting that eliminates them can be established.

Another limitation of this study concerns the age distributions of the TD (ages 3-10 years) and CP (ages 4 – 14 years) groups, which were different. Since the CP group included three children older than 10 years of age (the oldest age of children in the TD group), this limited the number of possible CP and TD matched pairs that could be identified for the discriminative validity analysis of accelerometry-based measures. As well, the slightly older mean age of the CP group compared to the TD group limits the extent to which these group’s criterion validity results are directly comparable to one another. Specifically, the values found for the accuracy (and sensitivity and specificity) of accelerometry-based measures with the TD group cannot be considered a direct norm reference for the values found for the clinical CP group in this study.

As previously described, there are psychometric limitations with some of the clinical measures used here. Both the PMAL-R and DDI do not have established validity based on a one-time
administration, as was done here. As well, the QUEST was not validated for the older child participants in this study (from 9 – 14 years of age), nor was the PMAL-R (for participants > 6 years of age).

5.9 Implications of Research Findings:

This study comprises the first step in the validation of a paediatric measurement tool that could be used in research and clinical settings to measure change and document dosage of therapy. Once further validation studies are conducted in more natural environments and other psychometric properties are investigated, arm-accelerometry could contribute to more objective measuring of targeted movements during therapeutic interventions for hemiparesis. As well, it could contribute to more objective measuring of treatment outcomes in children’s natural environments, making follow-up on the maintenance of targeted goals (of increased arm use) more accessible to clinicians and researchers. The information captured by arm accelerometers in children could in turn inform the evaluation and improvement of current and new therapeutic interventions for paediatric hemiparesis.
Chapter 6

6 Conclusions

6.1 Concluding Summary

Upper extremity interventions for paediatric hemiparesis often aim to increase children’s use of the hemiparetic arm in daily life. However, objective and accessible methods for clinicians and researchers to measure this construct post-treatment are lacking. Arm-accelerometry, which has been validated as an upper extremity outcome measure in post-stroke adults, was investigated here, based on its potential for use with paediatric populations with hemiparesis. As an initial validity test, accelerometry’s criterion concurrent, convergent and discriminative validity were examined, in a controlled lab setting, with a sample of typically-developing children and a group of children with hemiplegic CP.

Firstly, results indicated that arm accelerometry has criterion validity as a measure of paediatric arm movements, based on high agreement found between accelerometry-based scores and a criterion of video-based observations of arm movements.

Secondly, arm accelerometry was found to have convergent validity, based on a moderate correlation between accelerometry-based arm ratios and unilateral QUEST scores (Dissociated Movement) of a sample of eleven children with CP. This significant correlation was found when a high threshold setting was used that excluded very small arm movements. This result suggests that use of this higher threshold warrants further investigation, as it may allow for accelerometer measurements in children to better capture arm movements that are more related to arm use and function. No other significant correlations were found between arm ratios and other clinical upper extremity scores of children with CP tested here. However, this study’s power was limited by a small sample size.

Thirdly, arm accelerometry was shown to have discriminative validity based on significant differences found in arm ratios between matched pairs of children with and without CP. This result suggests that future research, investigating accelerometry’s potential to make finer
discriminations within groups of children with CP could be worthwhile. Specifically, differences could be tested in arm ratios between groups of children with different levels of arm impairment.

Collectively, the results obtained here indicate that arm accelerometry shows promise as a potential contributor to outcome measurement in clinical paediatric populations with hemiparesis. Further investigation of its psychometric properties with larger samples of children with CP, and in natural environments, is warranted. Based on the results of these future investigations, arm accelerometry could be useful in measuring both targeted arm movements during treatment sessions and changes in post-treatment arm use in children’s natural environments.
Chapter 7

7 Future Directions

7.1 Directions for Future Research

Based on the findings of this study, many future investigations of arm accelerometry in children with hemiparesis are recommended. Firstly, ongoing investigation of accelerometry’s convergent and discriminative validity, with larger samples of children with hemiplegic CP is recommended. These investigations could involve a re-running of the correlational analyses between arm ratios and QUEST, MM scores, and other clinical measures (i.e. DDI scores) with larger samples. As well, with larger samples of children at different levels of arm impairment, their arm ratios could be compared as an extended test of discriminative validity in children with CP.

Following the ongoing performance of lab-based tests of validity, the collection of accelerometry-based arm ratio data in children’s natural environments, over longer periods of time (i.e. 2-3 days) is recommended. This research could establish referent norms for arm ratios in TD and CP groups of children of different age groups.

It is further recommended that a criterion validity test be performed, investigating the accuracy, sensitivity and specificity at a higher threshold for movement detection that eliminates very small movements. Based on the results obtained here, there is potential for arm accelerometry’s value (i.e. convergent validity) to be established with use of this higher threshold. Thus, its accuracy at this higher threshold should be investigated and confirmed via use of another objective method (e.g. with 3D motion capture or Kinect-based video data) of measuring and excluding very small movements.

Other further research could explore the development of software programs for extracting arm ratios directly from accelerometry output, for use by clinicians. Such research and development would likely be necessary to increase the feasibility of use of this measurement tool in clinical practice.

Finally, testing of accelerometry-based ratios as measures of targeted affected arm use during treatment sessions for children with hemiparesis is recommended as a longer-term research
objective. As well, psychometric evaluations of accelerometry’s responsiveness to change could lead to this tool’s use as a post-treatment measure of change. Following these evaluations, accelerometry could be valuable in potentially assessing the efficacy of therapeutic interventions for children with hemiparesis.
References


Appendices
Appendix A

Data Collecton Sheet for all participants

Date: ______________________

Researcher: ______________________

Study ID#: ______________________

Child’s month/year of birth: ______________________

Male/female: ______________________

**Additional questions for parent/guardian of child with hemiparesis:**

Diagnosis: ______________ Month/year of post-natal stroke (if applicable): ______________

Side of weaker arm (left or right): ______________________

Is your child currently receiving any therapy (i.e. physical or occupational therapy or other)?

If yes, what type of therapy? And how often does your child attend this therapy?

MACS Level: ______________________

**Video Observation Notes:**

<table>
<thead>
<tr>
<th>5-min Block:</th>
<th>Notes:</th>
</tr>
</thead>
<tbody>
<tr>
<td>Unimanual activities</td>
<td></td>
</tr>
<tr>
<td>Bimanual activities</td>
<td></td>
</tr>
<tr>
<td>Free play</td>
<td></td>
</tr>
</tbody>
</table>
Additional clinical measures for children with hemiparesis:

Modified Tardieu Ratings (for affected upper extremity):

<table>
<thead>
<tr>
<th>R1 TARDIEU SPASTIC CATCH TEST ITEMS</th>
<th>R1 Spastic Catch Angle</th>
<th>Not Done/Not Tested</th>
<th>Instructions</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wrist Flexors</td>
<td>Right</td>
<td></td>
<td>Stretch wrist quickly from full wrist flexion to extension and estimate the angle of spastic catch with 0 representing full wrist flexion and 180 degrees representing full wrist extension</td>
</tr>
<tr>
<td></td>
<td>Left</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Elbow Flexors</td>
<td>Right</td>
<td></td>
<td>Stretch elbow quickly from full elbow flexion to extension and estimate the angle of spastic catch with 0 representing full elbow flexion and 180 degrees representing full elbow extension</td>
</tr>
<tr>
<td></td>
<td>Left</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Screening Test for Mirror Movements:

FINE MOTOR COORDINATION

(Start with non-hemiplegic hand)

<table>
<thead>
<tr>
<th>TEST ITEMS</th>
<th>Normal</th>
<th>Abnormal</th>
<th>Not Done/Tested</th>
<th>Mirror Movements Score of opposite *</th>
<th>Guidelines for Scoring</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rapid Sequential Finger Movements</td>
<td>Right</td>
<td></td>
<td></td>
<td></td>
<td><strong>Demonstrate</strong>: thumb touches tip of individual fingers (moving from index to 5th finger) 5 times &quot;As fast as you can&quot;</td>
</tr>
<tr>
<td></td>
<td>Left</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rapid Index Finger Tap</td>
<td>Right</td>
<td></td>
<td></td>
<td></td>
<td><strong>Demonstrate</strong>: seated, finger taps table top or distal thumb of the same hand, &quot;As fast as you can&quot; for approximately 10 taps</td>
</tr>
<tr>
<td></td>
<td>Left</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Activity</td>
<td>Right</td>
<td>Left</td>
<td><strong>Demonstrate</strong>: seated, hands loosely fisted, and asked to rotate forearm/fist from supination to pronation one hand at a time. Complete approximately 10 movements</td>
<td></td>
<td></td>
</tr>
<tr>
<td>----------------------------------------------</td>
<td>----------------------------</td>
<td>----------------------------</td>
<td>----------------------------------------------------------------------------------------------------------------------------------</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rotation of Fist by alternating supination and rotation of forearm</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pincer Grasp</td>
<td></td>
<td></td>
<td>Encourage to pick up small 2–3 mm. ball of rolled up paper</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Finger To Nose Testing</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Heel To Shin Testing</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rapid Foot Tap</td>
<td></td>
<td></td>
<td><strong>Demonstrate</strong>: feet flat on floor, foot taps floor 20 &quot;As fast as you can&quot;</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sitting/ Standing Balance</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Mirror Movements Scoring Guide:*

Please NOTE: When asking child to do activities using the Right Hand, if mirror movements are seen in the LEFT hand, then apply a mirror movement score to the LEFT hand (using the scoring guide to the right) and vice versa.

<table>
<thead>
<tr>
<th>Movement</th>
<th>Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>None</td>
<td>0</td>
</tr>
<tr>
<td>Barely discernable repetitive movements</td>
<td>1</td>
</tr>
<tr>
<td>Slight mirror movements</td>
<td>2</td>
</tr>
<tr>
<td>Strong, sustained repetitive movements</td>
<td>3</td>
</tr>
<tr>
<td>Movements equal to that expected for the intended hand</td>
<td>4</td>
</tr>
</tbody>
</table>

Clinical Measure: Quality of Upper Extremity Skills Test (QUEST) Scores

<table>
<thead>
<tr>
<th>Summary Scores</th>
<th>Transferred from QUEST Manual</th>
</tr>
</thead>
</table>
### Clinical Measure based on Structured Parent Interview:

**Pediatric Motor Activity Log – Revised (PMAL-R)**

<table>
<thead>
<tr>
<th>Summary Scores</th>
<th>Transferred from Parent Data</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Collection Sheet (PMAL-R)</td>
</tr>
</tbody>
</table>

#### How Often:

#### How Well:

### Baseline Measure based of Child’s Developmental Disregard Index:

<table>
<thead>
<tr>
<th>Component Scores</th>
<th>Values</th>
</tr>
</thead>
<tbody>
<tr>
<td>QUEST score (%) for Dissociated Movements:</td>
<td>A.</td>
</tr>
<tr>
<td>PMAL-R score (How Often)/5 x 100%</td>
<td>B.</td>
</tr>
<tr>
<td>Disregard Index score (A. – B.)</td>
<td>C.</td>
</tr>
</tbody>
</table>

A. Dissociated Movements

B. Grasps

C. Weight Bearing

D. Protective Extension

Total Score = Sum of Scores/ Total # of Sections